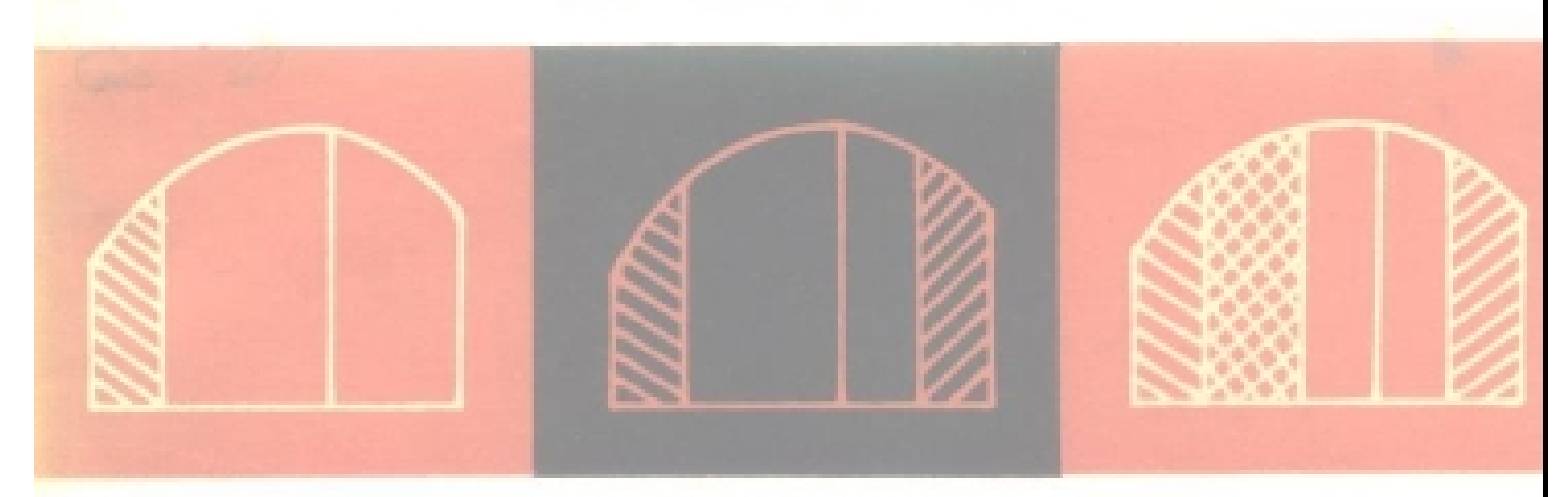
华罗庚 著





优选学

科学出版社

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

51.9132
673



优 选 学

华罗庚 著

1981.3.
图书馆藏

科 学 出 版 社
1 9 8 1

1109191

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

D68:1/17.

内 容 简 介

本书介绍优选学方面的一些常见的方法．全书共分三部分．第一部分单因素优选法，主要是介绍黄金分割法、分数法和抛物线法．第二部分多因素优选法，主要是介绍双因素优选法、最陡上升法、切块法、二次回归法和抛物体法．第三部分附录，介绍其它一些方法．

本书可供高等学校数学系师生以及运筹学工作者参考使用．

优 选 学

华罗庚著

*

科学出版社出版

北京朝阳门内大街137号

中国科学院印刷厂印刷

新华书店北京发行所发行　各地新华书店经售

*

1981年4月第一版　开本：787×1092 1/32

1981年4月第一次印刷　印张：5 7/8

印数：0001—14,750　字数：130,000

统一书号：13031・1513

本社书号：2076・13—1

定价：0.92元

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

序　　言

这本书的原稿完成于六十年代末期，作为向国庆二十周年的献礼，但一直未能出版，经过了十年，终于在国庆三十周年问世了．

优选法的试点、推广工作，曾经在毛主席和周总理的关怀下，在全国各地广泛地开展起来．

优选法是尽可能少做试验、尽快地找到生产最优方案的方法，也就是现代科学试验的方法．过去总是夸耀某人花了几十年功夫，做了几百次、甚至几千次的试验，获得了某一个科学成果．自从摆脱了以往排列组合做试验而出现了优选法之后，我们就要反问一句，是否有必要做那么多次试验，花那么长的时间才能获得某一成果．当然，花几十年搞试验的精神和毅力是令人钦佩的．可是，用优选法搞几十年就可以得到更多更好的成果．

优选法首先应用在化工、电子工业，然后逐步推广到各行各业．连炸油条省油，做豆腐省豆都用上了．在医学上，外科医生用了优选法能够在病人肠子里找到出血点．

为了使广大群众能够掌握和运用这个方法，曾经编写了几本通俗易懂的小册子．例如，《优选法平话》(署名齐念一，科学出版社，1971 年)、《优选法平话及其补充》(国防工业出版社，1971 年)、《优选法话本》(辽宁人民出版社，1973 年)．

近年来，我和各省市组织的优选法推广队伍跑遍了祖国二十二个省、市、自治区，到一百多个大小城镇，向广大群众介绍这个方法，并且推广各地组织的优选法经验．这个科学方

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

法已经被千百万工人所直接掌握．他们应用这个工具，获得了数以百万计的优选法成果（这里指的是推广的成果数而不是发明创造的成果数）．在国家不增加投资、设备、人力的条件下，为实现优质、高产、低消耗，作出了十分喜人的贡献．

这些年来，我们受到了极大的教育和启发，提高了对“群众性”和“实践性”的认识．我们体会到，一个方法如果只能被少数人所掌握，那末它所创造的劳动价值是很有限的．采用通俗易懂的平话形式，向广大群众介绍这个科学方法，是实现“群众性”和“实践性”的关键．但也有少数人产生了误解，认为平话没有什么理论．其实不然，从六十年代以来，我们就学习了几十种优选法，首先进行了理论上的推敲、剖析和研究，然后去粗取精，去伪存真，最后才挑选出几种理论上靠得住、又容易推广的和有效的方法．选定之后，我们又把这些方法深入浅出地通俗化，这样才能被广大群众所掌握．实践证明：在理论上深入固不易，而深入后浅出更不易．没有理论上的深入就没有方法上的浅出．

我们认为，一个科学成果从理论上研究到把它普及到群众中去，其中差距是不小的。后者涉及到千百万人的实践，因此，必须采取十分严谨的态度，必须经过反复实践．

这本《优选学》提到了不少不适宜于普及的东西．例如，关于多因素问题的处理，可能不下于国外七十年代所提到的一些方法，在理论上也可能更严密些．但是，我们还是应当以“抓主要矛盾”的思想为指导，在错综复杂的因素繁多的问题中，尽可能地抓住形成主要矛盾方面的因素，并且因素的个数也尽可能少．

现在我们伟大的祖国正进入社会主义现代化建设的新的历史时期，提高整个中华民族的科学文化水平是实现新时期总任务所必要的，从这几年广大群众掌握和应用优选法过程

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

中促进了学习技术、提高业务的情况，反映了这一科学方法在群众中的传播与推广，可以为实现这一宏伟目标作出一些贡献．从运用优选、统筹所创造的大量价值中，也有力地显示出科学就是生产力这一马克思主义原理的正确性．

在近年来的实践中，对如何把科学方法变成广大群众手中的工具，对如何洋为中用，对如何组织群众性的科学实验活动，以及如何发展适合于我国国情的科学技术等一系列问题，对我个人来说，总算与从前偈促于书斋之中，受囿于专题之下，不从国家需要和科学技术的相互关系考虑问题有所不同了．尽管自己在认识上提高了一大步，但是，这种认识还是粗浅的．

这本书虽然推迟了十年才出版，但从学术水平来说，还未曾落后．为了把全党工作着重点转移到实现四个现代化上来，为庆祝国庆三十周年，谨以此书聊表心意，献给伟大的祖国，献给伟大的人民．

最后，对参加本书校阅工作的龚昇、吴方、潘承烈、谢庭藩、徐伟宣、陈德泉、那吉生、裴定一、计雷、李之杰、方伟武等同志，以及在工作实践中予以协助的推广优选法小分队的同志们，表示衷心的感谢．同时希望读者和参加过这一工作的同志继续提出宝贵意见，以便进一步修改补充．

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

目　　录

第一部分　单因素优选法

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

第二部分 多因素优选法

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

第三部分 附录

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

第一部分　单因素优选法

我们常假定 $f(x)$ 是定义在区间 (a, b) 的单峰函数，但 $f(x)$ 的表达式是并不知道的，只有从试验中才能得出在某一点 x_0 的数值 $y_0 = f(x_0)$. 我们的问题是：用尽量少的试验次数来确定 $f(x)$ 的最大值的近似位置.

最简单的方法是穷举法或称均分法，即在 n 个点

$$x_l = a + \frac{l}{n+1}(b-a) \quad (l = 1, 2, \cdots, n)$$

做试验得 $y_l = f(x_l)$，其中最大的一个是 y_m，则我们知道最大值在区间

$$\left[a + \frac{m-1}{n+1}(b-a),\ a + \frac{m+1}{n+1}(b-a)\right]$$

中取，也就是做 n 次试验可以确定最大值在长度等于

$$\frac{2}{n+1}(b-a)$$

的区间中取，如果预先给了精密度 ε，则试验次数 n 必须适合于

$$\frac{2}{n+1}(b-a) < \varepsilon$$

才能达到目的，即：

$$n + 1 > \frac{2(b-a)}{\varepsilon}$$

所需要做的试验次数的数量级是$\frac{1}{\varepsilon}$才能达到目的.

以后所讲的黄金分割法及分数法，所需要的次数的数量

1109191

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

级只要 $\log\frac{1}{\varepsilon}$，大大改进了以上的结果，而抛物线法所需要的次数的数量级只要 $\log\log\frac{1}{\varepsilon}$，又进一步改进了黄金分割法的结果。

就试验的次数来说，方法一个精似一个，但这几个方法都是有序贯性的，即必须利用上一次试验的结果，才能决定下一次做试验的地点，这样做可以节省试验的次数，节省人力、物力及试验费用消耗。但是，可能要拉长时间，如果条件许可（例如有多套试验设备），我们可以用第三章所介绍的方法——一批做几个试验的方法。

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

第一章　黄金分割法和分数法

述　　要

《优选法平话》一书介绍过 0.618 法. 0.618 是

$$w=\frac{\sqrt{5}-1}{2}=0.6180339887\cdots$$

的四位有效近似值，它适合于方程

$$w^2+w-1=0.$$

黄金分割法也就是先在

$$x_1=a+(b-a)w$$

处做一次试验，再在 x_1 的对称点

$$x_2=b-(b-a)w=a+(b-a)w^2$$

做一次试验，比较试验结果 $y_1=f(x_1)$ 及 $y_2=f(x_2)$ 哪个大. 如果 $f(x_1)$ 大，就去掉 (a, x_2)，在留下的范围 (x_2, b) 中已有了一个试验点 x_1，然后再用以上的求对称点的方法做下去. 如果 $f(x_2)$ 大，则去掉 (x_1, b)，在留下的范围 (a, x_1) 中已有一试验点 x_2，同样用求对称点的方法做下去，一直做到所需要的精密度为止.

在介绍分数法之前，引进数列

$$F_0=1,\ F_1=1,\ F_n=F_{n-1}+F_{n-2}(n\geqslant 2),$$

即为

$$1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, 144, \cdots.$$

第 n 级的分数法以

$$\frac{F_n}{F_{n+1}}$$

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

代替黄金分割法的 $w=\frac{\sqrt{5}-1}{2}$，做了 n 次试验后，知道最大值在长度为 $\frac{2}{F_{n+1}}$ 的区间中，因此，给了精密度 ε，取 n 使

$$F_{n+1}\geqslant\frac{2}{\varepsilon}>F_n,$$

则用第 n 级的分数法做 n 次试验就可以达到要求。

§1. 有限点的问题——分数法

不妨假定所考虑的区间就是 $[0,1]$，在 m 个点 x_i

$$0<x_1<x_2<x_3<\cdots<x_m<1$$

处可以做试验。问题：至少要做几次试验可以找到 $f(x_i)$ $(1\leqslant i\leqslant m)$ 中最大的一个？或者我们至少要做几次试验，怎样安排，可以从最多的点数中找到取最大值的点来？

令 $m=\Phi_n=\Phi(n)$ 是 n 次试验可分辨的最多点数，也就是在 $m=\Phi(n)$ 个点中，可以做 n 次试验找到一个 x_i，使 $y_i=f(x_i)$ $(1\leqslant i\leqslant m)$ 为最大。而超过 Φ_n 个点，我们就无法用 n 次试验找出取最大值的点了。

首先，做一次试验 $(n=1)$，只能分辨一个点，因此 $\Phi_1=1$。做二次试验 $(n=2)$，只能在两个点中分辨，因此 $\Phi_2=2$。

考虑一般的情况：在 $\Phi_n=\Phi(n)$ 个点中已做了两次试验①，②如下图：

$$\underbrace{\circ\ \circ\ \cdots\ \circ}_{a}\ ①\ \underbrace{\circ\ \circ\ \cdots\ \circ}_{b}\ ②\ \underbrace{\circ\ \circ\ \cdots\ \circ}_{c}$$

即在 a 个点之后做一次试验①，再在 b 个点之后做一次试验②，还余 c 个点，则得

$$\Phi_n=a+b+c+2. \qquad (1)$$

如果①的试验结果比②好，则最好值一定在

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

$$\underbrace{\circ\ \circ\ \cdots\ \circ}_{a}\ ①\ \underbrace{\circ\ \circ\ \cdots\ \circ}_{b}$$

中，连①在内，在其中共可做 $n-1$ 次试验，可分辨的点数最多为 Φ_{n-1}，因此

$$a+b+1\leqslant\Phi_{n-1}. \tag{2}$$

又在前面 a 个点中最多只能再做 $n-2$ 次试验，所以

$$a\leqslant\Phi_{n-2}. \tag{3}$$

同样，如果②点比①点好，则应当有

$$b+c+1\leqslant\Phi_{n-1}. \tag{4}$$

综合 (1)，(2)，(3)，(4)，得

$$\Phi_n\leqslant\Phi_{n-1}+\Phi_{n-2}+1. \tag{5}$$

当且仅当 (2)，(3)，(4) 均为等式时等号成立。即当安排方法符合条件

$$a=c=\Phi_{n-2}\quad(\text{对称性}), \tag{6}$$

$$a+b+1=b+c+1=\Phi_{n-1}$$

时才能使可分辨的点数最多。这时

$$\Phi_n=\Phi_{n-1}+\Phi_{n-2}+1. \tag{7}$$

改写成递归公式

$$\Phi_1=1,\ \Phi_2=2,\ \Phi_n+1=(\Phi_{n-1}+1)+(\Phi_{n-2}+1). \tag{8}$$

令 $F_{n+1}=\Phi_n+1$，则得

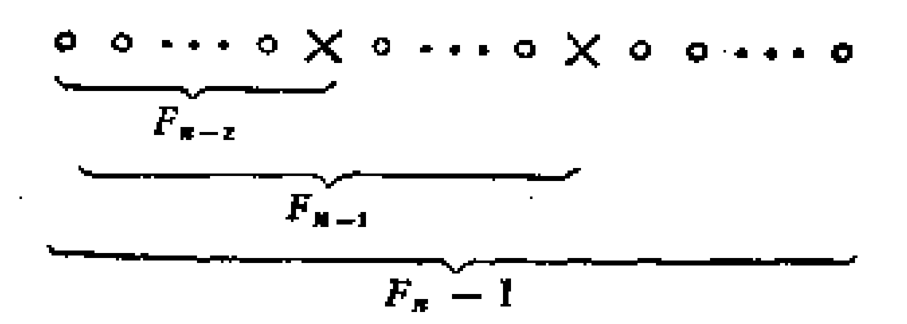
$$F_2=2,\ F_3=3,\ F_n=F_{n-1}+F_{n-2}\ (n=4,5,\cdots).$$

如果定义 $F_0=F_1=1$，上面的递归公式可写为

$$F_0=F_1=1,\ F_n=F_{n-1}+F_{n-2}\ (n=2,3,\cdots). \tag{9}$$

结论：做 n 次试验顶多可以分辨出 $(F_{n+1}-1)$ 个点中取最大值的那一点。安排方法是先在第 F_n 点处做试验，再在其对称点(即第 $F_{n+1}-F_n=F_{n-1}$ 点)处做试验，如果左边的试验点好，去掉第 F_n 点及以右各点，余下 (F_n-1) 个点，依法进行下去：

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn



如果这 $(F_{n+1}-1)$ 个点在 $(0,1)$ 中是均匀的，即均分 $(0,1)$ 为 F_{n+1} 等分，经过 n 次试验，最后余下长度为 $\frac{2}{F_{n+1}}$ 的区间，在 $(0,1)$ 中取最大值的点一定在这区间中。试验安排是，第一试验点在 F_n/F_{n+1} 处，以后取它的对称点，这就是第 n 级分数法。另一方面，做了 n 次试验以后，如果试验点恰好是余下

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

$$\left(\frac{1\pm\sqrt{5}}{2}\right)^{n}+\left(\frac{1\pm\sqrt{5}}{2}\right)^{n-1}$$

$$=\left(\frac{1\pm\sqrt{5}}{2}\right)^{n-1}\left(\frac{1\pm\sqrt{5}}{2}+1\right)$$

$$=\left(\frac{1\pm\sqrt{5}}{2}\right)^{n-1}\left(\frac{3\pm\sqrt{5}}{2}\right)$$

$$=\left(\frac{1\pm\sqrt{5}}{2}\right)^{n+1},$$

便有
$$F_n=F_{n-1}+F_{n-2}.$$
由归纳法证明了公式(1).

这里的证明是假定了已知公式(1)，再用归纳法来证明的。如果不知道公式(1)，如何导出这公式?

令

$$f(x)=F_0+F_1x+F_2x^2+\cdots+F_nx^n+\cdots, \tag{2}$$

$(2)\times(-x)$: $-xf(x)=-F_0x-F_1x^2-\cdots-F_{n-1}x^n-\cdots,$

$(2)\times(-x^2)$: $-x^2f(x)=\qquad -F_0x^2-\cdots-F_{n-2}x^n-\cdots,$

三式相加，得

$$(1-x-x^2)f(x)=F_0+(F_1-F_0)x$$
$$+\sum_{n=2}(F_n-F_{n-1}-F_{n-2})x^n=1,$$

即
$$f(x)=\frac{1}{1-x-x^2}. \tag{3}$$

令 $\quad 1-x-x^2=(1-w_1x)(1-w_2x),$

容易算出

$$w_1=\frac{1+\sqrt{5}}{2},\quad w_2=\frac{1-\sqrt{5}}{2}, \tag{4}$$

于是

$$f(x)=\frac{1}{(1-w_1x)(1-w_2x)}=\frac{A_1}{1-w_1x}+\frac{A_2}{1-w_2x}$$

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

$$= A_1\sum_{n=0}^{\infty}(w_1x)^n + A_2\sum_{n=0}^{\infty}(w_2x)^n$$

$$= \sum_{n=0}^{\infty}(A_1w_1^n + A_2w_2^n)x^n, \tag{5}$$

其中

$$A_1 = \lim_{x\to\frac{1}{w_1}} f(x)(1 - w_1x) = \lim_{x\to\frac{1}{w_1}}\frac{1}{1 - w_2x} = \frac{w_1}{w_1 - w_2},$$

同样

$$A_2 = -\frac{w_2}{w_1 - w_2}.$$

(5)式可写为

$$f(x) = \sum_{n=0}^{\infty}\frac{1}{w_1 - w_2}(w_1^{n+1} - w_2^{n+1})x^n, \tag{6}$$

与(2)式比较系数，便得

$$F_n = \frac{1}{w_1 - w_2}(w_1^{n+1} - w_2^{n+1}), \tag{7}$$

即为(1)式.

这个等式告诉我们，虽然所有的 F_n 都是正整数，可是它们可以用一些无理数表示出来. 由于 $\frac{\sqrt{5}-1}{2} < 1$，因此，当 n 是奇数时，则 F_n 等于 $\frac{1}{\sqrt{5}}\left(\frac{1+\sqrt{5}}{2}\right)^{n+1}$ 的整数部份，当 n 是偶数时，则 F_n 等于$\frac{1}{\sqrt{5}}\left(\frac{1+\sqrt{5}}{2}\right)^{n+1}$ 的整数部份加 1.

读者可参阅《从杨辉三角谈起》(青年数学小丛书，中国青年出版社).

作为练习，请读者求出下面两个数贯的解析表达式:

1) 数贯 $F_n^{(t)}$，由递归公式

$F_0^{(t)} = 1,\ F_1^{(t)} = t,\ F_n^{(t)} = t(F_{n-1}^{(t)} + F_{n-2}^{(t)})\ (n \geq 2)$

给出. 当 $t = 1$ 时，就是 F_n.

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

2）数贯 c_n，由递归公式

$$c_0 = c_1 = c_2 = 1, \ c_{n+1} = c_{n-1} + c_{n-2} \ (n \geqslant 2)$$

给出。

它们将分别在第二章和第三章中用到。

§3. 黄金分割法

分数法虽好，但预先需要知道做多少次试验，这是很难的，因为预先给定精密度（误差范围）ε 是不容易的。一般说来，函数的变化不是均匀的，越靠近极大值，变化越平坦，一旦试验结果不能分辨优劣时，试验就只好停止。所以，我们希望能有一个预先不限定试验次数的好方法。

现在我们来看看，当 n 趋向于无穷时，$\dfrac{F_n}{F_{n+1}}$ 的极限如何。

因为 $\left|\dfrac{1-\sqrt{5}}{2}\right| < 1$，易见

$$\lim_{n\to\infty} \frac{F_n}{F_{n+1}} = \lim_{n\to\infty} \frac{\left(\dfrac{1+\sqrt{5}}{2}\right)^{n+1} - \left(\dfrac{1-\sqrt{5}}{2}\right)^{n+1}}{\left(\dfrac{1+\sqrt{5}}{2}\right)^{n+2} - \left(\dfrac{1-\sqrt{5}}{2}\right)^{n+2}}$$

$$= \frac{\sqrt{5}-1}{2}.$$

我们记 $\omega = \dfrac{\sqrt{5}-1}{2}$.

进一步考虑差

$$\frac{F_n}{F_{n+1}} - \omega = \frac{F_n - F_{n+1}\omega}{F_{n+1}}$$

$$= \frac{1}{\sqrt{5}\,F_{n+1}} \left\{-\left(\frac{1-\sqrt{5}}{2}\right)^{n+1} - \left(\frac{1-\sqrt{5}}{2}\right)^{n+3}\right\}$$

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

$$=\frac{1}{F_{n+1}}\left(\frac{1-\sqrt{5}}{2}\right)^{n+2}=\frac{(-1)^n}{\left(\frac{1+\sqrt{5}}{2}\right)^{n+2}F_{n+1}}.$$

$$=\frac{-(-1)^{n-1}}{\left(\frac{1+\sqrt{5}}{2}\right)^{n+1}F_n}\cdot\frac{1}{\frac{1+\sqrt{5}}{2}\cdot\frac{F_{n+1}}{F_n}}$$

$$=\frac{w-\frac{F_{n-1}}{F_n}}{\left(\frac{1+\sqrt{5}}{2}\right)\frac{F_{n+1}}{F_n}},$$

$$\therefore\quad\left|\left(\frac{F_n}{F_{n+1}}-w\right)\Big/\left(w-\frac{F_{n-1}}{F_n}\right)\right|=\frac{1}{\frac{\sqrt{5}+1}{2}\cdot\frac{F_{n+1}}{F_n}}<1.$$

因此

$$\left|\frac{F_n}{F_{n+1}}-w\right|<\left|w-\frac{F_{n-1}}{F_n}\right|.$$

而由

$$\frac{F_n}{F_{n+1}}-w=\frac{(-1)^n}{\left(\frac{1+\sqrt{5}}{2}\right)^{n+2}F_{n+1}},$$

易见：当 n 为偶数时，$\frac{F_n}{F_{n+1}}>w$，当 n 为奇数时，$\frac{F_n}{F_{n+1}}<w$.

于是

$$\frac{F_2}{F_3}>\frac{F_4}{F_5}>\frac{F_6}{F_7}>\cdots>w,$$

$$\frac{F_1}{F_2}<\frac{F_3}{F_4}<\frac{F_5}{F_6}<\cdots<w.$$

这说明 $\frac{F_n}{F_{n+1}}$，当 n 趋于无穷时，在 w 的一左一右摆动，而且一个比一个更接近 w.

用黄金分割法如下：

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

再在 $(x_1\ b)$ 内取一点 x_3，做试验得 $y_3=f(x_3)$，如果 $x_3>x_2$ 而 $y_3<y_2$，则丢掉 (x_3, b)，再在 (x_1, x_3) 中取一点 x_4，···，不断做下去，就这样通过来回调试，范围越缩越小，总可以找到 $f(x)$ 的最大值(图 1)。

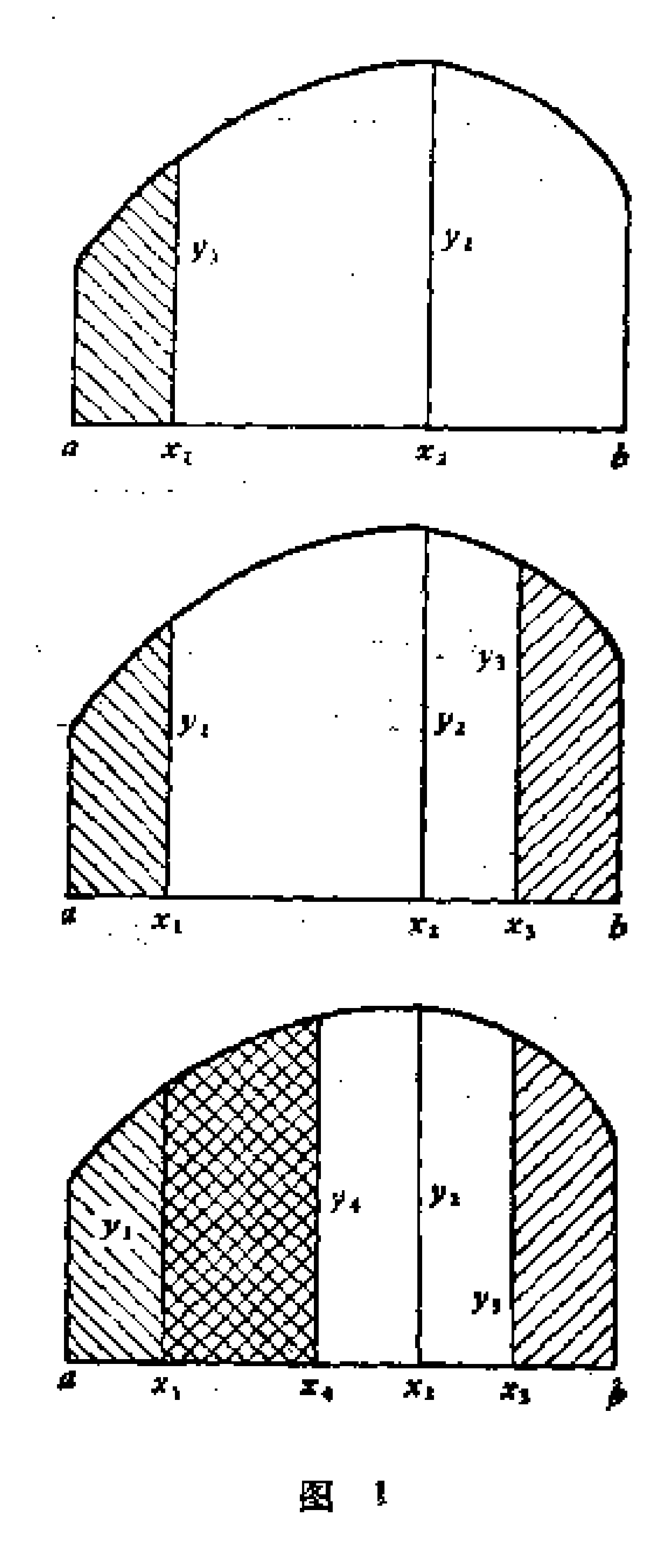


图 1

这方法取点是相当任意的，只要取在上次剩下的范围内就行了，我们的根本问题是：怎样取 x_1，x_2，···，可以最快地接近客观上存在的最高点？也就是怎样的安排试验点的方法是最好的？

何谓"最好"的来回调试法？先看一例，有一方法，第一试验点在 2/10，第二试验点在 1/10，对峰值在 (0，1/10) 中的单峰函数，两次试验便去掉了区间长度的 4/5 (图 2)，但对于峰值在 (2/10，1) 的函数，只能去掉 1/10，这就吃亏了。 由于我们试验的目标函数是事先全然不知道的，因此我们说的"最好"，就不能对个别的单峰函数来说，而应对全体这类函数而言。

这里不引入更多的数学定义，只是根据上面所说的事实，在评判一个来回调试法的效率时，作如下的约定：

如图 3 在 x_1，$x_2(x_1<x_2)$ 做了两个试验以后进行比较时，

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

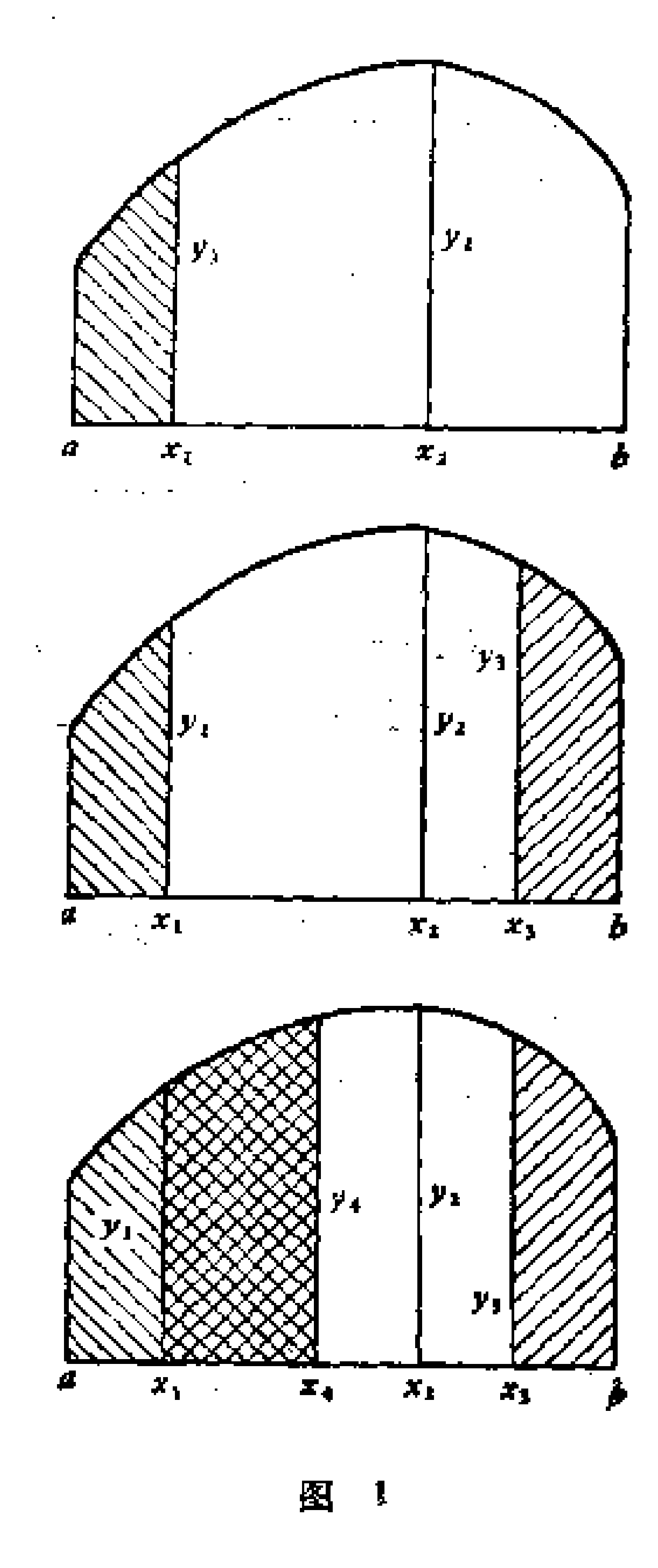


图 1

再在 $(x_1\ b)$ 内取一点 x_3，做试验得 $y_3=f(x_3)$，如果 $x_3>x_2$ 而 $y_3<y_2$，则丢掉 (x_3, b)，再在 (x_1, x_3) 中取一点 x_4，···，不断做下去，就这样通过来回调试，范围越缩越小，总可以找到 $f(x)$ 的最大值(图 1)。

这方法取点是相当任意的，只要取在上次剩下的范围内就行了。我们的根本问题是：怎样取 x_1，x_2，···，可以最快地接近客观上存在的最高点？也就是怎样的安排试验点的方法是最好的？

何谓"最好"的来回调试法？先看一例，有一方法，第一试验点在 2/10，第二试验点在 1/10，对峰值在 (0，1/10) 中的单峰函数，两次试验便去掉了区间长度的 4/5（图 2）。但对于峰值在(2/10，1)的函数，只能去掉 1/10，这就吃亏了。由于我们试验的目标函数是事先全然不知道的，因此我们说的"最好"，就不能对个别的单峰函数来说，而应对全体这类函数而言。

这里不引入更多的数学定义，只是根据上面所说的事实，在评判一个来回调试法的效率时，作如下的约定：

如图 3 在 x_1，$x_2(x_1<x_2)$ 做了两个试验以后进行比较时，

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

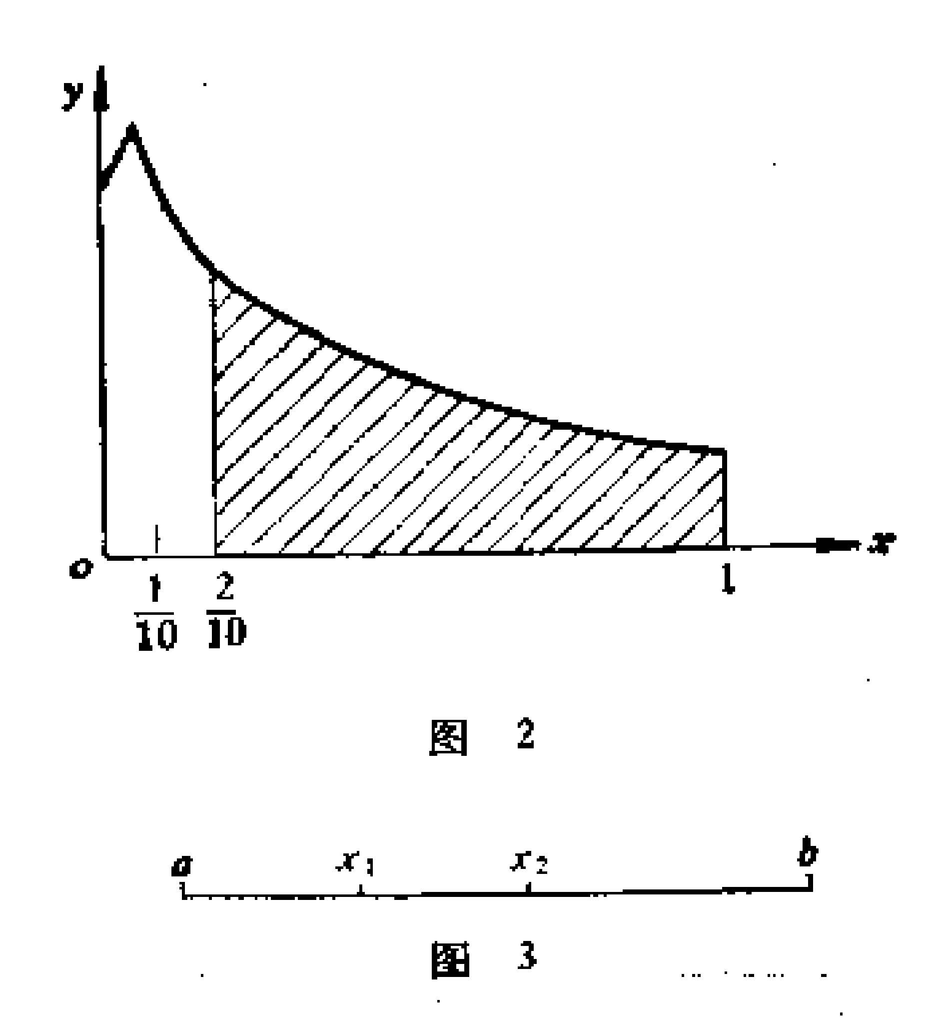


图 2

a　　x_1　　x_2　　b

图 3

总是遇到坏(吃亏)的情况，即只能去掉(a, x_1)和(x_2, b)中较短的那一段。

这样做了n次之后，余下一个包含有最好点的区间，其长度称为(该方法)n次试验的Δ-精密度。

对黄金分割法，假定原区间长度为1，则做n次试验后余下区间长度为w^{n-1}，于是

$$\Delta_n = w^{n-1}.$$

因此，若预先给定精密度ε，这时所要做的试验次数n应适合

$$w^{n-1} < \varepsilon,$$

即

$$n > 4.8 \log \frac{1}{\varepsilon} + 1.$$

也就是说n的数量级是$\log \frac{1}{\varepsilon}$.

在所有的来回调试法中，以黄金分割法为“最好”。我们有

基本定理　任何一个来回调试法，如果不是黄金分割法，总可以找到正整数N，当试验次数n大于N之后，其精度Δ_n

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

总不如黄金分割法做同样次数 n 的精度 ln，即

$$\Delta_n > \ln = \omega^{n-1} \quad (n > N).$$

§5. 黄金分割法的最优性

在这一节里，我们将证明上节的基本定理.

先建立一个不等式.

对于任意一个来回调试法，假定第一点和第二点分别在 x_1 和 x_2 处做，不妨假定区间就是 (0，1)，且 $x_1 < x_2$，显然 $\Delta_1 = 1$（图 4）.

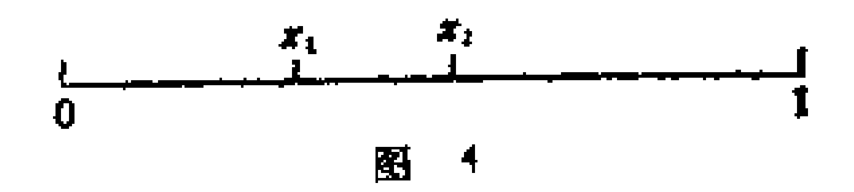


图 4

当 $x_1 + x_2 \geqslant 1$ 时，$\Delta_2 \geqslant x_2$，不论第三点如何取，总有

$$\Delta_3 \geqslant x_1,$$

故得不等式

$$\Delta_2 \leqslant \Delta_1 \leqslant \Delta_2 + \Delta_3. \tag{1}$$

当 $x_1 + x_2 < 1$ 时，$\Delta_2 \geqslant 1 - x_1$，

不论第三点如何取，总有 $\Delta_3 \geqslant 1 - x_2$，

故

$$\Delta_2 + \Delta_3 \geqslant (1 - x_1) + (1 - x_2) = 2 - (x_1 + x_2) > 1 = \Delta_1.$$

于是，不论第一、二、三点如何取，(1) 式恒成立.

一般地，我们有下面的不等式，

$$\Delta_{K+1} \leqslant \Delta_K \leqslant \Delta_{K+1} + \Delta_{K+2} \quad (K = 1, 2, \cdots). \tag{2}$$

由 (2)，用归纳法容易证明

$$\Delta_K \leqslant F_n \Delta_{n+K} + F_{n-1} \Delta_{n+K+1} \quad (K, n = 1, 2, \cdots). \tag{3}$$

这里 $F_0 = F_1 = 1$，$F_n = F_{n-1} + F_{n-2}$ $(n \geqslant 2)$.

记黄金分割法第 K 次试验后余下区间长度为 l_K，已知

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

$$l_K = w^{K-1}, \quad w = \frac{\sqrt{5}-1}{2} \tag{4}$$

对于 l_K，恒有等式

$$l_K = l_{K+1} + l_{K+2} \quad (K = 1, 2, \cdots), \tag{5}$$

$$l_K = F_n l_{n+K} + F_{n-1} l_{n+K+1} \quad (n, K = 1, 2, \cdots). \tag{6}$$

下面，用反证法来证明基本定理.

若定理不成立，则对于给定的来回调试法，必有一无穷数贯 $n_1 < n_2 < \cdots < n_s < \cdots$，

使得

$$\Delta_{n_s} \leqslant l_{n_s} \quad (s = 1, 2, \cdots). \tag{7}$$

对任意的 K，必有 n_i，使 $K < n_i < n_{i+1} < \cdots$. 不妨就当作 $n_0 = K < n_1 < n_2 < \cdots$.

由不等式(3)，得

$$\begin{aligned}
\Delta_K = \Delta_{n_0} &\leqslant F_{n_1-n_0}\Delta_{n_1} + F_{n_1-n_0-1}\Delta_{n_1+1} \\
&\leqslant F_{n_1-n_0}\Delta_{n_1} + F_{n_1-n_0-1}(F_{n_2-n_1-1}\Delta_{n_2} + F_{n_2-n_1-2}\Delta_{n_2+1}) \\
&\leqslant \cdots \\
&\leqslant F_{n_1-n_0}\Delta_{n_1} + F_{n_1-n_0-1}F_{n_2-n_1-1}\Delta_{n_2} + \cdots \\
&\quad + \left(\prod_{i=0}^{s-1} F_{n_{i+1}-n_i-1}\right)\Delta_{n_s} \\
&\quad + \left(\prod_{i=0}^{s-2} F_{n_{i+1}-n_i-1}\right)F_{n_s-n_{s-1}-2}\Delta_{n_s+1}.
\end{aligned} \tag{8}$$

而由(6)，有

$$\begin{aligned}
l_K = l_{n_0} &= F_{n_1-n_0}l_{n_1} + F_{n_1-n_0-1}F_{n_2-n_1-1}l_{n_2} + \cdots \\
&\quad + \left(\prod_{i=0}^{s-1} F_{n_{i+1}-n_i-1}\right) l_{n_s} \\
&\quad + \left(\prod_{i=0}^{s-2} F_{n_{i+1}-n_i-1}\right) F_{n_s-n_{s-1}-2} l_{n_s+1}.
\end{aligned} \tag{9}$$

由于 $\qquad \Delta_{n_i} \leqslant l_{n_i}, \quad \Delta_{n_s+1} \leqslant \Delta_{n_s} \leqslant l_{n_s},$

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

从(8)式减去(9)式，得

$$\Delta_K - l_K \leqslant \left(\prod_{i=0}^{s-2} F_{n_{i+1}-n_i-1}\right) F_{n_s-n_{s-1}-2}(l_{n_s} - l_{n_s+1})$$

$$= \left(\prod_{i=0}^{s-2} F_{n_{i+1}-n_i-1}\right) F_{n_s-n_{s-1}-2} l_{n_s+2}. \qquad (10)$$

因为

$$F_n = \frac{1}{\sqrt{5}}(w^{-(n+1)} - (-w)^{n+1})$$

$$\leqslant \frac{w^{-n}}{\sqrt{5}}\left(\frac{1}{w} + w\right) = w^{-n},$$

由(10)即得

$$\Delta_K - l_K \leqslant w^{s-1-n_{s-1}+n_0} \cdot w^{-n_s+n_{s-1}+2} \cdot w^{n_s+1}$$

$$= w^{s+n_0+2} = w^{s+K+2} \quad (s = 1, 2, \cdots).$$

令 $s \to \infty$，即得

$$\Delta_K \leqslant l_K \quad (K = 1, 2, \cdots). \qquad (11)$$

由定理假定，给定的来回调试法不是黄金分割法，因而必有某一个 s，使

$$\Delta_s < l_s.$$

由(3)和(6)，

$$1 = \Delta_1 \leqslant F_{s-1}\Delta_s + F_{s-2}\Delta_{s+1}$$

$$< F_{s-1}l_s + F_{s-2}l_{s+1} = l_1 = 1,$$

得出矛盾.

证毕.

§6. 连分数的知识

现在，我们换一个方法研究来回调试法，给出黄金分割法最优性的一个直接证明，并说明一些更深刻的内容.

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

为此，需要用到连分数这一工具。如果我们约定只考虑对称取点的来回调试法，那末就与辗转相除法、连分数关系十分密切，抽象地说，他们实际上是一回事。所以，我们先介绍连分数的知识。

1) 先从两个正整数 a, b 谈起。假定 $b < a$，用 b 除 a 得商 a_0，余 r，

$$a = a_0 b + r, \quad 0 \leqslant r < b.$$

若 $r \neq 0$，再用 r 除 b 得商 a_1，余 r_1，即

$$b = a_1 r + r_1, \quad 0 \leqslant r_1 < r;$$

若 $r_1 \neq 0$，再用 r_1 除 r，得

$$r = a_2 r_1 + r_2, \quad 0 \leqslant r_2 < r;$$

若 $r_2 \neq 0$，则继续做下去，由此得

$$b > r > r_1 > r_2 > \cdots \geqslant 0$$

经有限步后，余数必为 0，例如 $r_n = 0$，这个方法叫做辗转相除法。上面的过程可以写为：

$$\frac{a}{b} = a_0 + \frac{r}{b} = a_0 + \cfrac{1}{a_1 + \cfrac{r_1}{r}} = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{r_2}{r_1}}}$$

$$= \cdots = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{\ddots + \cfrac{1}{a_n}}}},$$

这一繁分数称为连分数。为了节省篇幅，把它写为

$$\frac{a}{b} = a_0 + \frac{1}{a_1 +}\,\frac{1}{a_2 +} \cdots \frac{1}{+ a_n}.$$

例 1

$$\frac{1}{2} = \frac{1}{1 +}\,\frac{1}{1}, \qquad \frac{2}{3} = \frac{1}{1 +}\,\frac{1}{1 +}\,\frac{1}{1},$$

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

$$\frac{3}{5}=\frac{1}{1}+\frac{1}{1}+\frac{1}{1}+\frac{1}{1},$$

一般地有

$$\frac{F_n}{F_{n+1}}=\frac{1}{1+\frac{F_{n-1}}{F_n}}=\underbrace{\frac{1}{1}+\frac{1}{1}+\cdots+\frac{1}{1}}_{n+1\text{个}}.$$

2）辗转相除不一定限于有理数 $\frac{a}{b}$，对任一无理数 α，也可以写为

$$\alpha=a_0+\frac{1}{\alpha_1},\quad a_0=[\alpha],\quad \alpha_1>1,$$

方括弧代表 α 的整数部分．再令

$$\alpha_1=a_1+\frac{1}{\alpha_2},\quad a_1=[\alpha_1],\quad \alpha_2>1.$$

一般地

$$\alpha_{n-1}=a_{n-1}+\frac{1}{\alpha_n},\quad a_{n-1}=[\alpha_{n-1}],\quad \alpha_n>1.$$

因此，得到 α 的连分数表达式为

$$a_0+\frac{1}{a_1}+\frac{1}{a_2}+\cdots+\frac{1}{a_{n-1}}+\cdots.$$

因为，任意有限项的连分数都是有理数，所以无理数的连分数必然是无穷的．

经计算得

$$a_0+\frac{1}{a_1}=\frac{a_0a_1+1}{a_1},$$

$$a_0+\frac{1}{a_1}+\frac{1}{a_2}=\frac{a_0a_1a_2+a_2+a_0}{a_2a_1+1},$$

……

令

$$a_0+\frac{1}{a_1}+\frac{1}{a_2}+\cdots+\frac{1}{a_n}=\frac{p_n}{q_n},$$

称为 α 的第 n 个渐近分数．

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

例 2 令 $w=\frac{\sqrt{5}-1}{2}$，它适合于 $w^2+w-1=0$，

即 $w=\frac{1}{1+w}=\frac{1}{1}+\frac{1}{1+w}=\frac{1}{1}+\frac{1}{1}+\frac{1}{1+w}$

$$=\frac{1}{1}+\frac{1}{1}+\frac{1}{1}+\cdots$$

因此，$\frac{F_n}{F_{n+1}}$ 就是 w 的渐近分数，§3 中的结果说明

$$\frac{F_2}{F_3}>\frac{F_4}{F_5}>\frac{F_6}{F_7}>\cdots>w,\quad \frac{F_1}{F_2}<\frac{F_3}{F_4}<\frac{F_5}{F_6}<\cdots<w.$$

即得

$$\frac{F_{2i}}{F_{2i+1}}>w>\frac{F_{2i-1}}{F_{2i}}$$

的一个小于一个的区间套，又由

$$\begin{aligned}\frac{F_{2i}}{F_{2i+1}}-\frac{F_{2i-1}}{F_{2i}}&=\frac{1}{5F_{2i}F_{2i+1}}\{[w^{-(2i+1)}-(-w)^{2i+1}]^2\\&\qquad-[w^{-2i}-(-w)^{2i}][w^{-2i-2}-(-w)^{2i+2}]\}\\&=\frac{1}{5F_{2i}\cdot F_{2i+1}}(w^2+w^{-2}+2)\\&=\frac{1}{5F_{2i}\cdot F_{2i+1}}[w^{-1}-(-w)]^2=\frac{1}{F_{2i+1}F_{2i}},\end{aligned}$$

可知区间套一个套着一个。

§7. 连分数与来回调试法（一）

考虑对称取点的来回调试法。

在区间 $(0,1)$ 中已做了试验的点为 α。

(i) 若 $\alpha=\frac{1}{1+\theta_1}$，$0<\theta_1<1$。

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

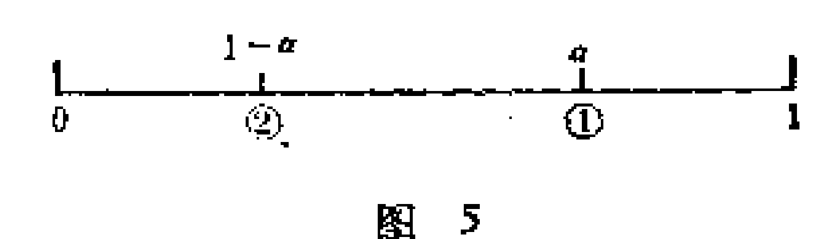


图 5

则 $\alpha > \frac{1}{2}$，对称取点，第二个试验点在 $1-\alpha$ 处，不论去掉那一段，余下区间的长度总是 α，

即

$$\Delta_1 = 1, \quad \Delta_2 = \alpha = \frac{1}{1+\theta_1}.$$

不失一般性，不妨假定总是靠近左边的点好，也就是说，总是去掉右边一小段。

因为 $\alpha + \theta_1\alpha = 1$，即 $1-\alpha = \theta_1\alpha = \theta_1\Delta_2$，因而，从 α 开始做二次试验后，余下区间中的已试验点，即 $(1-\alpha)$ 的位置恰好在区间长度的 θ_1 倍处，同样，若

$$\theta_1 = \frac{1}{1+\theta_2}, \quad 0 < \theta_2 < 1.$$

则用对称法取第三点，余下的区间长度为

$$\Delta_3 = \theta_1\Delta_2 = \theta_1\alpha.$$

余下区间中的已试验点的位置在区间长度的 θ_2 倍处。

(ii) 若 $\alpha = \frac{1}{p+\theta}$，$p$ 为 $\geqslant 2$ 的整数，$0 < \theta < 1$，

0　α　1－α　1

图 6

则 $\alpha < \frac{1}{2}$，从 α 开始，对称取点做二次试验后，余下区间的长度为

$$\Delta_2 = 1 - \alpha = 1 - \frac{1}{p+\theta} = \frac{p-1+\theta}{p+\theta}$$

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

$$= \cfrac{1}{1 + \cfrac{1}{p - 1 + \theta}}, \tag{1}$$

其中已试验点(即 α)的位置在区间长度的

$$\alpha:(1-\alpha) = \frac{1}{p-1+\theta}$$

倍处，再做下去相当于在区间中从 $\alpha_1 = \dfrac{1}{p-1+\theta}$ 出发来回调试，再做了 $p-2$ 次后，余下区间 $(0, 1-(p-1)\alpha)$，即由一段长为 $\theta\alpha$，另一段为 α 的二段所组成，所以再做一次，取点应在 $\theta\alpha$ 处了，余下区间的长度为 α，其中已试验点的位置恰好在区间长度的 θ 倍处。

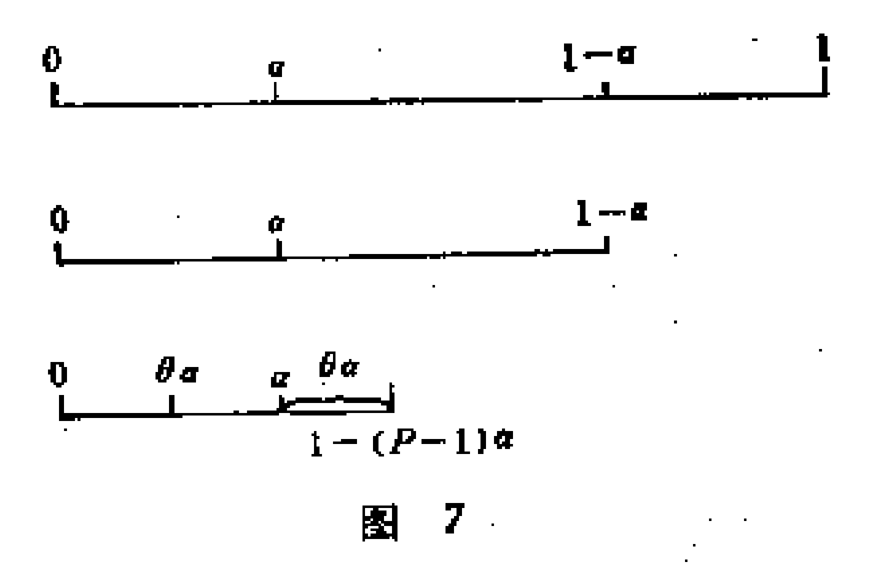


图 7

这样不断做下去，如

$$\alpha = \frac{1}{a_1} + \frac{1}{a_2} + \cdots + \frac{1}{a_n + \theta_n}, \quad 0 < \theta_n < 1.$$

则经 $a_1 + a_2 + \cdots + a_n$ 次对称取点调试后(不算 α 本身一次)，余下区间中的已试验点的位置恰好在长度的 θ_n 倍处。

再证明几个引理。

引理 1 若

$$\alpha = \underbrace{\frac{1}{1} + \frac{1}{1} + \cdots + \frac{1}{1 + \theta}}_{m\text{个}}, \quad 0 < \theta < 1.$$

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

则

$$\alpha=\frac{F_{m-1}+\theta F_{m-2}}{F_m+\theta F_{m-1}},\tag{2}$$

这里 $F_{-1}=0$，$F_0=1$，$F_n=F_{n-1}+F_{n-2}(n\geqslant 1)$.

证　用数学归纳法.

当 $m=1$ 时，$\alpha=\frac{1}{1+\theta}$，(2) 显然成立.

当 $m=2$ 时，

$$\alpha=\frac{1}{1+\frac{1}{1+\theta}}=\frac{1+\theta}{2+\theta},$$

(2) 也成立.

当 $m=n$ 时，令 $\theta_{n-1}=\frac{1}{1+\theta}$，

则由归纳法假定，即得

$$\alpha=\frac{F_{n-2}+\theta_{n-1}F_{n-3}}{F_{n-1}+\theta_{n-1}F_{n-2}}=\frac{F_{n-2}+\frac{1}{1+\theta}F_{n-3}}{F_{n-1}+\frac{1}{1+\theta}F_{n-2}}$$

$$=\frac{F_{n-1}+\theta F_{n-2}}{F_n+\theta F_{n-1}}.$$

故 (2) 式成立.

引理 2　不等式

$$\frac{1}{F_{m+2}}\geqslant\frac{1}{5w^2}\cdot w^m\quad(m\geqslant 1)$$

成立，当且仅当 $m=2$ 等号成立，而

$$\frac{1}{5w^2}=\frac{3+\sqrt{5}}{10}=0.523\cdots.$$

证　仍用归纳法.

当 $m=1$ 时，$\frac{1}{F_3}=\frac{1}{3}=\frac{1}{3w}w>\frac{1}{5w^2}\cdot w$.

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

当 $m=2$ 时，$\frac{1}{F_4}=\frac{1}{5}=\frac{1}{5w^2}\cdot w^2$.

故当 $m=1$, 2 时，引理 2 成立.

对一般的 m，若 $\frac{F_{m+1}}{F_{m+2}}>w$，则由归纳法假设，

$$\frac{1}{F_{m+2}}=\frac{F_{m+1}}{F_{m+2}}\cdot\frac{1}{F_{m+1}}>w\left(\frac{1}{5w^2}\right)w^{m-1}=\frac{1}{5w^2}\cdot w^m.$$

若 $\frac{F_{m+1}}{F_{m+2}}<w$，则由 § 3，$\frac{F_m}{F_{m+1}}>w$，

即 $\frac{F_{m+1}}{F_m}<\frac{1}{w}$，所以由归纳法假设，

$$\frac{1}{F_{m+2}}=\frac{1}{F_m+F_{m+1}}=\frac{1}{F_m}\cdot\frac{1}{1+\frac{F_{m+1}}{F_m}}$$

$$>\left(\frac{1}{5w^2}\right)w^{m-2}\frac{1}{1+\frac{1}{w}}$$

$$=\left(\frac{1}{5w^2}\right)w^{m-2}\cdot\frac{w}{1+w}=\frac{1}{5w^2}\cdot w^m.$$

故对一般 m 引理 2 也成立.

同样可证

引理 3 不等式

$$\frac{1}{F_{m+1}}\geqslant\frac{1}{2w}w^m\quad(m\geqslant1)$$

成立，当且仅当 $m=1$ 时等号成立，而

$$\frac{1}{2w}=\frac{1}{\sqrt{5}-1}=0.745\cdots.$$

现在来证 § 4 的基本定理.

设区间 (0, 1) 中已有试验点 α.

1. 若

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

$$\alpha=\underbrace{\frac{1}{1}+\frac{1}{1}+\cdots+\frac{1}{1+\theta_m}}_{m\text{ 个}},\quad 0<\theta_m<1,$$

记

$$\alpha=\frac{1}{1+\theta_1}=\frac{1}{1+\frac{1}{1+\theta_2}}=\cdots=\frac{1}{1}+\cdots+\frac{1}{1+\theta_m},$$

则再做 m 次(共 $m+1$ 次)试验后，由引理 1 可知，余下区间长度为

$$\Delta_{m+1}=\alpha\theta_1\theta_2\cdots\theta_{m-1}=\frac{F_{m-1}+\theta_m F_{m-2}}{F_m+\theta_m F_{m-1}}\cdot\frac{F_{m-2}+\theta_m F_{m-3}}{F_{m-1}+\theta_m F_{m-2}}\cdot\cdots\cdot\frac{1}{1+\theta_m}$$

$$=\frac{1}{F_m+\theta_m F_{m-1}}.$$

因为 $0<\theta_m<1$，故由引理 3，可得

$$\Delta_{m+1}>\frac{1}{F_m+F_{m-1}}=\frac{1}{F_{m+1}}\geqslant\frac{1}{2w}w^m\quad(m\geqslant 1).$$

II. 若

$$\alpha=\underbrace{\frac{1}{1}+\frac{1}{1}+\cdots+\frac{1}{1+\theta_m}}_{m\text{ 个}},$$

其中 $\theta_m=\dfrac{1}{p+\theta_{m+1}},\quad m\geqslant 1,\quad p\geqslant 2,\quad 0<\theta_{m+1}<1.$

由于 $\theta_m<\dfrac{1}{2}$，则由引理 2 可得

$$\Delta_{m+1}=\frac{1}{F_m+\theta_m F_{m-1}}>\frac{1}{F_m+\frac{1}{2}F_{m-1}}=\frac{2}{2F_m+F_{m-1}}$$

$$=\frac{2}{F_{m+2}}\geqslant\frac{2}{5w^2}w^m\quad(m\geqslant 1).$$

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

III. 若

$$\alpha=\frac{1}{p+\theta},\quad p\text{ 为 }\geqslant 2\text{ 的整数},\quad 0<\theta<1,$$

分二种情形来讨论.

i）若 $p\geqslant 3$，则由(1)，再做一次试验后，

$$\Delta_2=\frac{1}{1+\frac{1}{p-1+\theta}}>\frac{1}{1+\frac{1}{2}}=\frac{2}{3}=\frac{2}{3\omega}\omega.$$

ii）若 $p=2$，(0，1)中已试验点在 $\alpha=\frac{1}{2+\theta}$ 处，第二个试验点应在 $1-\alpha$ 处，此时，$\alpha<\frac{1}{2}<1-\alpha$，这与(0，1)中已试验点在 $1-\alpha$，再做第二个试验点在 α 的情形是一样的。而

$$1-\alpha=1-\frac{1}{2+\theta}=\frac{1+\theta}{2+\theta}=\frac{1}{1+\frac{1}{1+\theta}},$$

归结为 I 的情形.

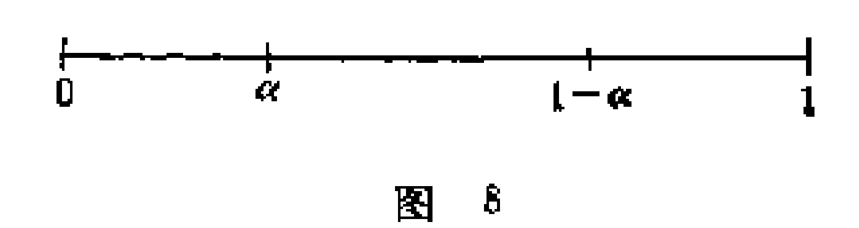


图 8

因之，若

$$\alpha=\frac{1}{1}+\cdots+\frac{1}{1}+\frac{1}{p+\theta_m},\quad p\geqslant 2,\quad 0<\theta_m<1.$$

在(0，1)中依 α 进行对称来回调试，由 II，做 $m+1$ 次试验后，余下区间中已试验点的位置在 $\frac{\Delta_{m+1}}{p+\theta_m}$，$p\geqslant 2$. 再由 III i)，再做 $p-2$ 次试验后，即共做 $m+1+p-2=m+p-1$ 次后，余下的区间长度为

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

$$\Delta_{m+p-1} \geqslant \left(\frac{2}{5w^2}\right) w^m \left(\frac{2}{3w}\right)^{p-2} w^{p-2} > \left(\frac{2}{5w^2}\right)^{p-1} w^{m+p-2}.$$

再往下做就是 III ii) 的情形了.

现在我们回到基本定理的证明.

若 α 是有理数，可由 § 8 的论证中顺便推得基本定理成立.

若 α 不是有理数，即 α 可以写成无穷连分数，此时，对称取点来回调试法可以一直做下去.

如 $\alpha \fallingdotseq w$，那末 α 的展开式中一定不全是1，记这些不为 1 的数依次为 $p_1, p_2, p_3, \cdots$.

1. 若这样的 p_i 多于 c 个（c 在下面定出），则当 n 充分大之后，综合 I, II, III, 恒有

$$\Delta_n \geqslant \left(\frac{2}{5w^2}\right)^{\sum_{i=1}^{c}(p_i-1)} \frac{1}{2w} w^{n-1}.$$

因子 $\frac{1}{2w} = 0.745\cdots < 1$，至多是在最后一个试验点是属于 I 或 III (ii) 的情形时才出现.

如果 $\left(\frac{2}{5w^2}\right)^{\sum_{i=1}^{c}(p_i-1)} > 2w$，即

$$\sum_{i=1}^{c}(p_i - 1) > \frac{\log 2w}{\log 2/5w^2} = 4.6\cdots.$$

时，则 $\Delta_n > \ln$，只要取 $c = 5$ 就可以达到目的，即 p_i 多于 5 个时，当 n 充分大时，

$$\Delta_n > w^{n-1} = \ln.$$

2. 若这样的 p_i 不超过 $c - 1$ 个，则

$$\alpha = \frac{1}{1} + \cdots \frac{1}{p_1} + \frac{1}{1} + \cdots + \frac{1}{p_{c-1}} + \frac{1}{1} + \frac{1}{1} + \cdots$$

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

$$= \frac{1}{1+}\cdots\frac{1}{p_1+}\frac{1}{1+}\cdots+\frac{1}{p_{c-1}+w}.$$

故做充分大 n 次试验以后，又变为黄金分割法了，此时

$$\Delta_n \geqslant \left(\frac{2}{5w^2}\right)^{\sum_{i=1}^{c-1}(p_i-1)} w^{n-1} > w^{n-1} = l_n.$$

这就证明了基本定理。

附记 衡量一个方法好坏时，还可以用另一种精度，即以 n 次试验后余下区间中已试验点与较远的那个端点的距离 δ_n 作为 n 次试验的精密度。如 n 次试验后余下的区间为 (a, b)，其中已试验点为 x，真正的最好点为 x_0，于是

$$|x_0 - x_n| \leqslant \delta_n = \max(x-a, b-x).$$

图 9

为了与前面所说的 Δ- 精密度区别，称这种精密度为 δ- 精密度。

分数法 F_n/F_{n+1} 的 δ- 精密度为 $\delta_n = \frac{1}{F_{n+1}}$，黄金分割法的 δ- 精密度为 $\delta_n = w^n$。由于对称取点，总有 $\delta_n = \Delta_{n+1}$。从上面证明即得：在 δ- 精密度下，基本定理依然成立。

§8. 连分数与来回调试法（二）

上面证明了在不限定试验次数的情形下黄金分割法的最优性。现在再来看看用来回调试法限定做 N 次试验的情形。

设

$$\alpha = \underbrace{\frac{1}{1+}\frac{1}{1+}\cdots+\frac{1}{1+\theta_m}}_{m\text{个}}$$

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

其中 $m+p+1<N$，$\theta_m=\dfrac{1}{p+\theta_{m+1}}$，整数 $p\geqslant 2$，$0<\theta_{m+1}<1$．作

$$\alpha'=\underbrace{\frac{1}{1}+\frac{1}{1}+\cdots+\frac{1}{1}}_{m+1\text{个}}+\frac{1}{(p-1)+\theta_{m+1}}.$$

从 α 出发，用来回调试法，做 $m+1$ 次试验后，余下区间的长度为

$$\Delta_{m+1}=\frac{1}{F_m+\theta_m F_{m-1}},$$

其中已试验点的位置在 $\theta_m\Delta_{m+1}$ 处，再做一次，因为 $0<\theta_m<\dfrac{1}{2}$，故有

$$\Delta_{m+2}=(1-\theta_m)\Delta_{m+1}=\frac{1-\theta_m}{F_m+\theta_m F_{m-1}}$$
$$=\frac{p-1+\theta_{m+1}}{(p-1+\theta_{m+1})F_m+F_{m+1}},$$

其中已试验点的位置在

$$\frac{1}{p-1+\theta_{m+1}}\Delta_{m+2}$$

处．

从 α' 出发，用来回调试法，做 $m+2$ 次试验后，余下区间的长度为

$$\Delta'_{m+2}=\frac{1}{F_{m+1}+\dfrac{1}{p-1+\theta_{m+1}}F_m}$$
$$=\frac{p-1+\theta_{m+1}}{F_m+(p-1+\theta_{m+1})F_{m+1}},$$

其中已试点的位置在

$$\frac{1}{p-1+\theta_{m+1}}\Delta'_{m+2}$$

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

处，由于

$$F_m+(p-1+\theta_{m+1})F_{m+1}-[(p-1+\theta_{m+1})F_m+F_{m+1}]$$
$$=(p-2+\theta_{m+1})F_{m-1}>0,$$

故

$$\Delta_{m+2}>\Delta'_{m+2}$$

且它们中间已试验点的位置是一样的，故当 $N\geqslant n\geqslant m+2$ 时，却有

$$\Delta_n>\Delta'_n,$$

或一样有

$$\delta_n>\delta'_n.$$

也就是说，对于所有形式的 α：

$$\alpha=\frac{1}{a_1}+\frac{1}{a_2}+\cdots+\frac{1}{a_l+\theta},$$

其中 $a_1,\cdots,a_l$ 为正整数，$a_1+a_2+\cdots+a_l+1=N$，

$$1\leqslant l\leqslant N-1,\quad 0<\theta<1.$$

从 $\bar{\alpha}=\underbrace{\frac{1}{1}+\frac{1}{1}+\cdots+\frac{1}{1+\theta}}_{N-1\text{个}},\quad 0<\theta<1$

出发，用来回调试法做 N 次试验所得的精度为最好.

由此可知，只做 N 次试验的最好方法只要从 α 中去找就可以了. 而从 $\bar{\alpha}$ 出发，用来回调试法做 N 次试验后余下的区间长度为

$$\Delta_n=\frac{1}{F_{N-1}+\theta F_{n-2}},$$

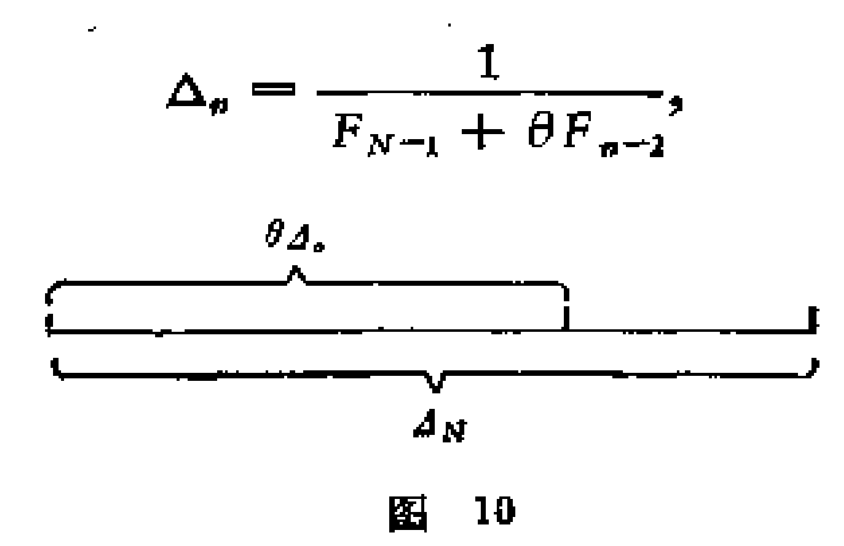


图 10

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

其中已试点的位置在 $\theta\Delta_N$ 处.

因而

$$\delta_N = \max(1-\theta, \theta)\Delta_N = \frac{\max(1-\theta, \theta)}{F_{N-1}+\theta F_{N-2}}$$

$$= \begin{cases} \dfrac{1-\theta}{F_{N-1}+\theta F_{N-2}}, & 若\ 0 < \theta \leqslant \dfrac{1}{2}. \\ \dfrac{\theta}{F_{N-1}+\theta F_{N-2}}, & 若\ \dfrac{1}{2} \leqslant \theta < 1. \end{cases}$$

显然，当 $\theta = \dfrac{1}{2}$ 时，δ_N 取最小值

$$\delta_N = \frac{\frac{1}{2}}{F_{N-1}+\frac{1}{2}F_{N-2}} = \frac{1}{F_{N+1}},$$

而此时，

$$\bar{\alpha} = \frac{1}{\underbrace{1+\frac{1}{1+\cdots+}}_{N-1个}\frac{1}{1+\frac{1}{2}}}$$

$$= \frac{1}{\underbrace{1+\frac{1}{1+\cdots+\frac{1}{1}}}_{N+1个}} = \frac{F_N}{F_{N+1}}.$$

这就是分数法.

这样，我们在 δ- 精度下，证明了分数法 $\dfrac{F_N}{F_{N+1}}$ 是 N 次试验的最优来回调试法，这与 § 1 中的结论是一致的.

显然

$$\delta_N\left(\frac{F_N}{F_{N+1}}\right) < \delta_N(\omega),$$

即

$$\frac{1}{F_{N+1}} < \omega^N.$$

由引理 3，可推出

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

$$\frac{1}{F_{N+1}} \geqslant \left(\frac{1}{2w}\right) w^N = \left(\frac{1}{2w^2}\right) w^{N+1} > w^{N+1} = \delta_{N+1}(w).$$

从而有

$$w^{N+1} < \frac{1}{F_{N+1}} < w^N.$$

即在 δ- 精度下，分数法比黄金分割法好．如用黄金分割法再多做一次试验就比分数法好了．

在 Δ- 精度下情况就不一样了． 由于 Δ_N 是 θ 在 $(0, 1)$ 中单调递减函数．当 $\theta \to 1$ 时，Δ_N 单调递减于

$$\frac{1}{F_{N-1} + F_{N-2}} = \frac{1}{F_N}$$

因为只限于做 N 次试验，故 θ 不能等于 1，也就是没有一个方法能达到精度 $\frac{1}{F_N}$，即不存在一个最好的方法．

分数法 $\frac{F_N}{F_{N+1}}$ 相当于 α 中 $\theta = \frac{1}{2}$ 的情形，而黄金分割法相当于 $\theta = w$ 的情形，由于 $w > \frac{1}{2}$，可知

$$\Delta_N(F_N/F_{N+1}) > \Delta_N(w),$$

即

$$\frac{2}{F_{N+1}} > w^{N-1}.$$

此式亦可从引理 3 直接推出，这说明在 Δ- 精度下，分数法又不如黄金分割法了．

到此，分数法与黄金分割法的关系就比较清楚了，并且解决了 α 是有理数的情形．

§9. 对 分 法

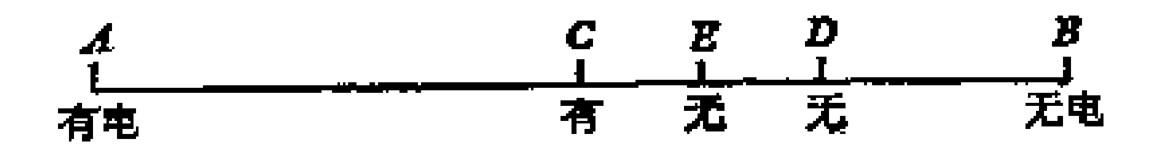
上面介绍的是：求函数 $f(x)$ 最优值的一些方法和它们

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

的理论根据．在实践中也常遇到这样一类特殊的问题，例如：在保证质量的前提下，如何降低某种原材料的消耗？如何降低成本？如何尽快地找到电路、通讯线路、下水道、输油管以及机器、仪器的故障？砂轮、转子的静平衡等．这些问题的特殊性在于：每次试验的结果只有二种性态，因此可以化为另一类的数学模型．

对这类问题，我们建议用以下的对分法，而不用 0.618 法．

例如，有一条 10 公里长的输电线路出了故障，在线路一端 A 处有电，在另一端 B 处没有电，要迅速查出断头所在位置．我们先在线路中间的 C 点检查，如果有电，说明断头不在 AC 而是在 CB 一段，又在 CB 中点 D 检查，如果无电，说明断头在 CD 部分，再在 CD 中间 E 处检查，如此类推，很快就能找出断头位置．



对分法最简单，但在实践中应用非常普遍．用对分法每做一次试验，区间长度缩短一半．若原区间长为 l，做了 n 次试验后，余下区间长度为

$$\frac{1}{2^n}l,$$

易见，对分法所需试验次数的数量级也是

$$\log\frac{1}{\varepsilon},$$

其中 ε 是预先给定的精度．

对一些不易求解的方程或方程组，用对分法能迅速求出近似解．可参阅本书附录二．

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

第二章　抛物线法

述　　要

不管是 0.618 法还是分数法，都只是比较两个试验结果的好坏，而不考虑已做试验结果的数值如何．能不能利用已得的试验数据，做到对技术精益求精呢？例如在试得三个数据后，我们可以过这三点作一抛物线，以这抛物线的达到顶点处作下次试验的根据．确切地说是这样：在三个点 $x_1 < x_2 < x_3$，分别试验得数据 y_1，y_2，y_3，我们用插入公式

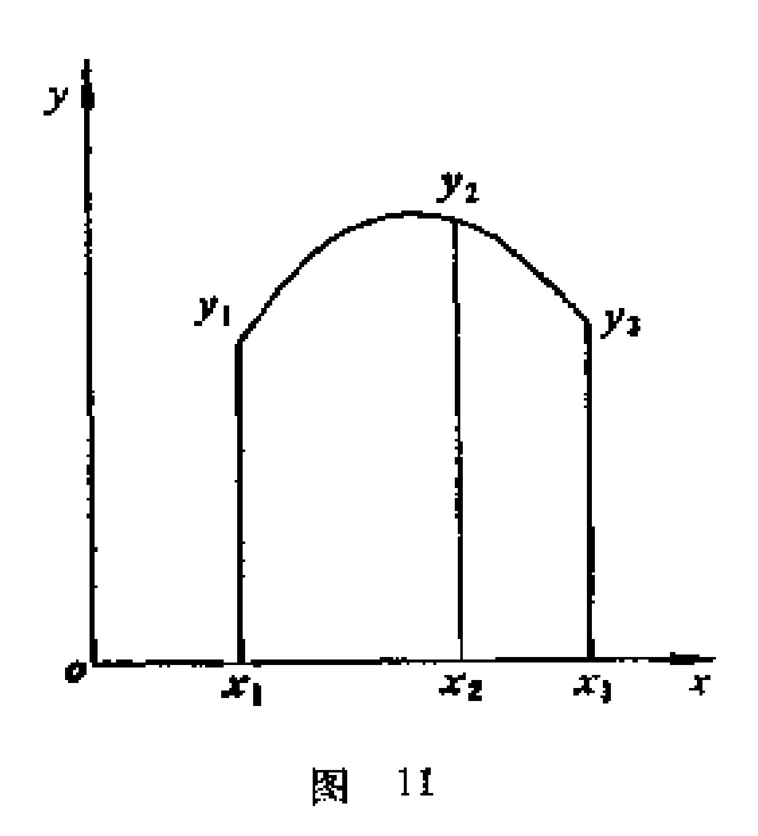


图　11

$$y = y_1 \frac{(x - x_2)(x - x_3)}{(x_1 - x_2)(x_1 - x_3)} + y_2 \frac{(x - x_3)(x - x_1)}{(x_2 - x_3)(x_2 - x_1)} + y_3 \frac{(x - x_1)(x - x_2)}{(x_3 - x_1)(x_3 - x_2)}$$

得到一个二次函数，它的图像就是一条抛物线．这个函数在点

$$x_4 = \frac{1}{2} \frac{y_1(x_2^2 - x_3^2) + y_2(x_3^2 - x_1^2) + y_3(x_1^2 - x_2^2)}{y_1(x_2 - x_3) + y_2(x_3 - x_1) + y_3(x_1 - x_2)}$$

处取最大值，因此，我们下一次的选点取 $x = x_4$，得试验数据 y_4，然后利用 x_2，x_3，x_4 三点的数据 y_2，y_3，y_4 又可得出一条抛物线，如此继续做下去．

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

当然，为了更快接近最大值，可在 y_1, y_2, y_3, y_4 中取最大的，所对应的 x_i 作为新的 x_2，并取其左边最近的 x_i 作为新的 x_1，右边最近的 x_i 作为新的 x_3，再同样做下去，不断获得新的 y 值.

如果发生 $x_4 = x_2$ 的情况，我们的方法还必须修改（§ 1）.

粗略地说，穷举法作 n 次试验的同样效果，黄金分割法只要数量级 $\log n$ 次就可以达到，本章提出的抛物线法，效果更好些，只要数量级 $\log\log n$ 次，原因就在于黄金分割法没有较多地利用函数的性质，做了两次试验，比一比大小，就把它丢掉了. 抛物线法则对试验结果注意到数量方面的分析. 当然，要保证抛物线法收敛速度，对函数也作了一些适当的限制.

抛物线法常常用在 0.618 法或分数法取得一些数据的情况，这时能收到更好的效果. 此外，我们还建议大家做完了 0.618 法或分数法的试验后，用最后三个数据按抛物线法求出 x_4，并计算这个抛物线在点 $x = x_4$ 处的数值，预先估计一下在点 x_4 处的试验结果，然后将这个数值与已经试得的（在 y_1, y_2, y_3 中）最佳值作比较，以此作为是否在点 x_4 处再作一试验的依据.

§ 1. 第一种方法

设在三点

$$x_1 < x_2 < x_3$$

处分别测得 $y = f(x)$ 的函数值

$$f(x_1) = y_1, \quad f(x_2) = y_2, \quad f(x_3) = y_3,$$

用 Lagrange 插入法作抛物线

$$g(x) = y_1 \frac{(x - x_2)(x - x_3)}{(x_1 - x_2)(x_1 - x_3)} + y_2 \frac{(x - x_3)(x - x_1)}{(x_2 - x_3)(x_2 - x_1)}$$

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

$$+ y_3 \frac{(x-x_1)(x-x_2)}{(x_3-x_1)(x_3-x_2)}.$$

这个函数的极值由

$$g'(x) \equiv y_1 \frac{2x - x_2 - x_3}{(x_1-x_2)(x_1-x_3)} + y_2 \frac{2x - x_3 - x_1}{(x_2-x_3)(x_2-x_1)}$$

$$+ y_3 \frac{2x - x_1 - x_2}{(x_3-x_1)(x_3-x_2)} = 0$$

给出，也就是

$$x = x_4 = \frac{1}{2} \frac{y_1(x_2^2 - x_3^2) + y_2(x_3^2 - x_1^2) + y_3(x_1^2 - x_2^2)}{y_1(x_2 - x_3) + y_2(x_3 - x_1) + y_3(x_1 - x_2)}$$

$$= \frac{1}{2} \frac{(x_3^2 - x_2^2)(y_2 - y_1) + (x_1^2 - x_2^2)(y_3 - y_2)}{(x_3 - x_2)(y_2 - y_1) + (x_1 - x_2)(y_3 - y_2)}$$

$$= x_2 + \frac{1}{2} \frac{(x_3 - x_2)^2(y_2 - y_1) + (x_1 - x_2)^2(y_3 - y_2)}{(x_3 - x_2)(y_2 - y_1) + (x_1 - x_2)(y_3 - y_2)}. \tag{1}$$

现在我们假定 $y = f(x)$，是向上凸的，也就是

$$y_2 = f(x_2) > \frac{(x_3 - x_2)f(x_1) + (x_2 - x_1)f(x_3)}{x_3 - x_1},$$

即

$$(x_3 - x_2)y_1 + (x_1 - x_3)y_2 + (x_2 - x_1)y_3 < 0. \tag{2}$$

这时候 $g''(x_4) < 0$，所以 x_4 正是 $g(x)$ 的最大值点.

于是，立刻建议出第一个方法：如果

$$y_2 > y_1, \quad y_2 > y_3,$$

那末，我们由公式（1）得到 x_4，在 $x = x_4$ 处做试验得函数 $y = f(x)$ 的值 y_4. 若 $y_4 > y_2$，取 x_4 作为新的 x_2，并取其左边最近的 x_i 作为新的 x_1，右边最近的 x_i 作为新的 x_3. 若 $y_4 \leqslant y_2$，仍取 x_2 作为新的 x_2，再同样做下去.

这样，我们就面临着两个问题：

1）这个方法做得下去吗？如果 $x_4 = x_2$，怎么办？

2）这个方法能否导致 $f(x)$ 的最大值？

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

先回答问题 2)，由 (1) 式及 (2) 式得到

$$x_4 = \frac{1}{2}(x_3 + x_2) - \frac{1}{2}\frac{(x_2 - x_1)(x_3 - x_1)(y_2 - y_3)}{(x_3 - x_2)(y_2 - y_1) + (x_2 - x_1)(y_2 - y_3)} < \frac{1}{2}(x_3 + x_2),$$

和

$$x_4 = \frac{1}{2}(x_1 + x_2) + \frac{1}{2}\frac{(x_3 - x_1)(x_3 - x_2)(y_2 - y_1)}{(x_3 - x_2)(y_2 - y_1) + (x_2 - x_1)(y_2 - y_3)} > \frac{1}{2}(x_1 + x_2).$$

假定

$$x_1 < x_2 < x_4 < x_3,$$

如果 $y_4 > y_2$，则函数 $f(x)$ 的最大值所在的区间由 (x_1, x_3) 缩成为 (x_2, x_3)；如果 $y_4 < y_2$，则缩成为 (x_1, x_4)，而

$$x_4 < \frac{1}{2}(x_3 + x_2).$$

同样处理 $x_4 < x_2$ 的情况.

总起来一句话：每做一次，区间缩短一次，从而得到一系列的前一个包着后一个的区间

$$(x_1^{(i)}, x_3^{(i)}), \quad (i = 1, 2, \cdots),$$

因而其极限存在，记作

$$(x_1^{(0)}, x_3^{(0)})$$

(请注意 $x_1^{(i)}$ 是不降的，$x_3^{(i)}$ 是不升的)，不难证明 $x_1^{(0)}$，$x_3^{(0)}$ 之一就是我们要找的点 x_0. 此处，我们不作证明的原因有二：①留给读者作练习，②下一节所谈的第二方法的证明比这个更

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

深刻.

现在回到问题 1)，当 $x_4 = x_2$ 时，可以看看 $x_3 - x_2$ 和 $x_2 - x_1$ 哪个大. 如果 $x_2 - x_1$ 较大,新的点 x_4' 可以按下法取.

(i) 取

$$x_4' = x_2 - \varepsilon.$$

这儿 $\varepsilon > 0$ 是要求的精密度,然后再用抛物线法做下去. 如果 $f(x_2)$ 与 $f(x_2 - \varepsilon)$ 不易分辨,则说明 x_2 可能已到顶了，换取 $x_4'' = x_2 + \varepsilon$. 如果此时还是 $f(x_2) > f(x_2 + \varepsilon)$，则下一试验点就取在 $x_2 + \varepsilon^2$ 的地方等等.

这样取点法优点是快,缺点是 ε 难以预估,试验结果可能难以分辨.

(ii) 取

$$x_4' = \frac{x_1 + x_2}{2},$$

$\left(若 x_3 - x_2 较大,则取 x_4' = \dfrac{x_2 + x_3}{2}\right)$ 这取法速度慢. 假如适逢我们的目标函数就是顶点在 x_2 的抛物线,那就吃亏了.

为了克服 (i) 和 (ii) 出现的缺点,可以采取

(iii) 在 (i) 中令

$$\varepsilon = k(x_3 - x_2)^2$$

其中 k 为定常数,要求其保证无论进行到哪一步都有

$$k(x_3 - x_2)^2 \leqslant \frac{1}{2}(x_3 - x_2),$$

可令 k 为原始区间长度二倍的倒数.

然后取 $x_4' = x_2 + \varepsilon$，再按抛物线法做下去. 这时显然有 $x_2 < x_4' \leqslant \dfrac{x_2 + x_3}{2}$. 这样，既照顾到分辨问题，又照顾到收敛速度同 $x_4 \neq x_2$ 的情况一致.

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

§2. 误差估计

假定函数 $f(x)$ 在 $x=x_0$ 处取最大值，抛物线法就是建议用上节公式（1）算出的 x_4 去近似地代替 x_0，本节将对误差 $|x_4-x_0|$ 作出估计。

显然，如果 x_0 不在区间端点，则

$$f'(x_0)=0. \tag{1}$$

由上节公式（1）可知

$$x_4-x_0=\left[\frac{1}{x_3-x_1}\left\{\frac{y_1-y_2}{x_1-x_2}-\frac{y_3-y_2}{x_3-x_2}\right\}(x_3+x_2-2x_0)\right.$$
$$\left.+\frac{y_3-y_2}{x_3-x_2}\right]\Big/\left[\frac{2}{x_3-x_1}\left(\frac{y_1-y_2}{x_1-x_2}-\frac{y_3-y_2}{x_3-x_2}\right)\right],$$

记其分子、分母分别为 N, D, 并令

$$x_i=x_0+\xi_i \quad (i=1,2,3,4).$$

则

$$N=\frac{1}{\xi_3-\xi_1}\left(\frac{y_1-y_2}{\xi_1-\xi_2}-\frac{y_3-y_2}{\xi_3-\xi_2}\right)(\xi_3+\xi_2)+\frac{y_3-y_2}{\xi_3-\xi_2},$$

$$D=\frac{2}{\xi_3-\xi_1}\left(\frac{y_1-y_2}{\xi_1-\xi_2}-\frac{y_3-y_2}{\xi_3-\xi_2}\right).$$

由于

$$y_i-y_j=f(x_0+\xi_i)-f(x_0+\xi_j)$$
$$=(\xi_i-\xi_j)\int_0^1 f'(x_0+\xi_j+(\xi_i-\xi_j)t)dt,$$

所以

$$D=\frac{2}{\xi_3-\xi_1}\int_0^1(f'(x_0+\xi_2(1-t)+\xi_1t)$$
$$-f'(x_0+\xi_2(1-t)+\xi_3t))dt$$
$$=-2\int_0^1\int_0^1 f''(x_0+\xi_2(1-t)+\xi_1t\mu$$
$$+\xi_3(1-\mu)t)td\mu dt.$$

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

如果
$$|f''(x)| \geqslant m_2, \tag{2}$$
则有
$$|D| \geqslant m_2. \tag{3}$$
事实上，我们在上节就已假定曲线 $y=f(x)$ 是向上凸的，因此 (2) 含有 $f''(x) \leqslant -m_2$，于是从 D 的表达式得到 (3)。

再考虑分子，类似于上述，有
$$N = -(\xi_3+\xi_2)\int_0^1\int_0^1 f''(x_0+\xi_2(1-t)+\xi_1 t\mu$$
$$+\xi_3(1-\mu)t)td\mu dt + \int_0^1 f'(x_0+\xi_2(1-t)$$
$$+\xi_3 t)dt.$$
但因 $f'(x_0)=0$，所以
$$\int_0^1 f'(x_0+\xi_2(1-t)+\xi_3 t)dt = \int_0^1 (f'(x_0+\xi_2(1-t)$$
$$+\xi_3 t)-f'(x_0))dt = \int_0^1\int_0^1 f''(x_0+\xi_2(1-t)\mu$$
$$+\xi_3 t\mu)(\xi_2(1-t)+\xi_3 t)d\mu dt,$$
及
$$\int_0^1\int_0^1 (\xi_2(1-t)+\xi_3 t)d\mu dt = \frac{1}{2}(\xi_2+\xi_3),$$
所以
$$N = -(\xi_3+\xi_2)\int_0^1\int_0^1 (f''(x_0+\xi_2(1-t)+\xi_1 t\mu$$
$$+\xi_3(1-\mu)t)-f''(x_0))td\mu dt$$
$$+\int_0^1\int_0^1 (f''(x_0+\xi_2(1-t)\mu+\xi_3 t\mu)$$
$$-f''(x_0))(\xi_2(1-t)+\xi_3 t)d\mu dt. \tag{4}$$
这样，如果 $|f'''(x)|$ 有界，即若
$$|f'''(x)| \leqslant k, \tag{5}$$
就有

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

$$|N| \leqslant |\xi_3 + \xi_2| \int_0^1\int_0^1 k|\xi_2(1-t) + \xi_1 t\mu + \xi_3(1 - \mu)t| t d\mu dt + \int_0^1\int_0^1 k|\xi_2(1-t)\mu + \xi_3 t\mu| |\xi_2(1-t) + \xi_3 t| d\mu dt$$

$$\leqslant k|\xi_3 + \xi_2|\left(\frac{|\xi_2|}{6} + \frac{|\xi_1|}{6} + \frac{|\xi_3|}{6}\right) + k\left(\frac{\xi_2^2}{6} + \frac{\xi_3^2}{6} + \frac{|\xi_3\xi_2|}{6}\right) \leqslant k\left(\frac{\xi_2^2}{3} + \frac{\xi_3^2}{3} + \frac{|\xi_2\xi_3|}{2} + \frac{|\xi_3\xi_1|}{6} + \frac{|\xi_2\xi_1|}{6}\right).$$

因此，我们得到以下结论：如果 $m_2 > 0$，

$$f''(x) \leqslant -m_2, \quad |f'''(x)| \leqslant k, \tag{6}$$

则有

$$|\xi_4| \leqslant \frac{3k}{2m_2}\max(|\xi_2|^2, |\xi_3|^2, |\xi_2\xi_3|, |\xi_3\xi_1|, |\xi_1\xi_2|), \tag{7}$$

§3. 第二种方法

根据上节最后一个不等式，我们可以求得另一方法，先用其他方法找到

$$x_1 < x_2 < x_3, \quad x_3 - x_1 < \frac{2m_2}{3k}q$$

及

$$y_2 > y_1, \quad y_2 > y_3.$$

这儿 $0 < q < 1$，因此

$$|\xi_i| < \frac{2m_2}{3k}q \quad (i = 1, 2, 3). \tag{1}$$

现在用以下的方法来确定试验点，当有了 x_{n-2}，x_{n-1}，和 x_n 之后，由

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

$$x_{n+1}=\frac{1}{2}[(x_{n-1}^2-x_n^2)y_{n-2}+(x_n^2-x_{n-2}^2)y_{n-1}$$
$$+(x_{n-2}^2-x_{n-1}^2)y_n]/[(x_{n-1}-x_n)y_{n-2}$$
$$+(x_n-x_{n-2})y_{n-1}+(x_{n-2}-x_{n-1})y_n]$$

定义 x_{n+1}，再测出 y_{n+1}，不管 y_{n-2}，y_{n-1}，y_n 和 y_{n+1} 中谁大谁小，我们把同一方法用在 x_{n-1}，x_n，x_{n+1} 上，下面证明这方法一定收敛，不仅如此，我们还要证明

$$|\xi_n|<\frac{2m_2}{3k}q^{c_n}, \tag{2}$$

这里 $c_1=c_2=c_3=1$，

$$c_n=c_{n-2}+c_{n-3}\quad(n=4,5\cdots). \tag{3}$$

由 § 2 不等式 (7) 得到

$$|\xi_n|\leqslant\frac{3k}{2m_2}\max(|\xi_{n-3}\xi_{n-2}|,\ |\xi_{n-2}\xi_{n-1}|,$$
$$|\xi_{n-1}\xi_{n-3}|,\ \xi_{n-2}^2,\ \xi_{n-1}^2) \tag{4}$$

因为 c_n 是一个不降的整数列，所以

$$\min(c_{n-3}+c_{n-2},\ c_{n-2}+c_{n-1},\ c_{n-1}+c_{n-3},$$
$$2c_{n-2},\ 2c_{n-1})=c_{n-3}+c_{n-2} \tag{5}$$

由 (1) 知道不等式 (2) 当 $n=1, 2, 3$ 时成立，为了完成 (2) 的证明，只要用数学归纳法就可以了，事实上，假设不等式 (2) 对 $n=1, 2, \cdots, k-1$ 都成立，那么由 (4)，(5) 两式得到

$$|\xi_k|<\frac{3k}{2m_2}\left(\frac{2m_2}{3k}\right)^2 q^{\min(c_{k-3}+c_{k-2},\ c_{k-2}+c_{k-1},\ c_{k-1}+c_{k-3},\ 2c_{k-2},\ 2c_{k-1})}$$
$$=\frac{2m_2}{3k}q^{c_{k-3}+c_{k-2}}.$$

即

$$|\xi_k|<\frac{2m_2}{3k}q^{c_k}.$$

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

§4. C_n 的表达式

现在我们研究适合于

$$c_{n+1}=c_{n-1}+c_{n-2},\quad c_0=c_1=c_2=1$$

的 c_n 的表达式. 先算出几项来:

1, 1, 1, 2, 2, 3, 4, 5, 7, 9, 12, 16, 21, 28, 37, 49, …,

求 c_n 的方法同第一章 §2 一样. 令

$$f(x)=\sum_{n=0}^{\infty}c_nx^n.$$

由

$$(1-x^2-x^3)\sum_{n=0}^{\infty}c_nx^n=1+x+x^2+\sum_{m=0}^{\infty}c_{m+3}x^{m+3}$$

$$-x^2-\sum_{m=0}^{\infty}c_{m+1}x^{m+3}-\sum_{m=0}^{\infty}c_mx^{m+3}=1+x,$$

得到

$$f(x)=\frac{1+x}{1-x^2-x^3}.$$

令

$$1-x^2-x^3=(1-w_1x)(1-w_2x)(1-w_3x),$$

即 w_i 是方程式

$$y^3-y-1=0$$

的根, 这方程有一个大于 1 的实根 $w_1=1.324\cdots$ 及两个绝对值小于 1 的共轭虚根.

用分项分数法,

$$f(x)=\frac{A_1}{1-w_1x}+\frac{A_2}{1-w_2x}+\frac{A_3}{1-w_3x},\qquad(1)$$

这儿

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

$$A_i = \lim_{x\to\frac{1}{w_i}}(1-w_ix)\frac{1+x}{1-x^2-x^3} = \frac{(1+x)w_i}{2x+3x^2}\bigg|_{x=\frac{1}{w_i}}$$

$$= \frac{(1+w_i)w_i^2}{2w_i+3}.$$

由 (1) 式得出

$$c_n = A_1w_1^n + A_2w_2^n + A_3w_3^n.$$

因为

$$\lim_{n\to\infty} w_2^n = \lim_{n\to\infty} w_3^n = 0$$

所以

$$c_n \sim A_1w_1^n$$

于是,我们有以下的收敛结果

$$|\xi_n| \leqslant cq^{w_1^n},$$

即 n 又升高了一个台阶.

附记 在考虑抛物线法的收敛速度时，如果把条件 $|f'''(x)| \leqslant k$ 换作

$$|f''(x'') - f''(x')| \leqslant k|x''-x'|$$

或者说 $f''(x) \in \mathrm{Lip}1$，自然是可以的. 或许有的读者会提出 $f''(x) \leqslant -m_2$，不易验证，那末把它换作初始三点 x_1, x_2, x 上 $f(x)$ 的二阶差分

$$\frac{2}{x_3-x_1}\left(\frac{y_1-y_2}{x_1-x_2} - \frac{y_3-y_2}{x_3-x_2}\right) \geqslant m_2 > 0 \qquad (2)$$

也是可以的,这就表明,对于给定的 f，开始用其他方法安排，一旦条件 (2) 适合时,再用抛物线法，收敛速度就更快了，例如,此时若 $x_3 - x_1 \leqslant \frac{2m_2}{3k}q$, $0 < q < 1$, 就易于证明

$$|x_n - x_0| \leqslant \frac{2m_2}{3k}q^{c_n}, \quad c_1 = c_2 = c_3 = 1,$$

$$c_n = c_{n-2} + c_{n-3} \quad (n = 4, 5, \cdots).$$

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

§5. 第三种方法

从三点 $x_1 < x_2 < x_3$ 出发，有了 $x_1, x_2, \cdots x_n$ 之后要去找 x_{n+1}，在具体应用抛物线法时，除前面讲的两个办法之外，我们再建议用如下的算法：

1）先从三点 $x_1 < x_2 < x_3$ 出发，并假定 $y_1 < y_2, y_3 < y_2$，即中间高，两头低，定义

$$a_3 = x_1, \quad b_3 = x_3, \quad c_3 = x_2, \quad a_3 < c_3 < b_3.$$

2）有了 $x_1, x_2, \cdots, x_n$ 之后，假设 $y_1, y_2, \cdots y_n$ 中最大的一个在 $x = c_n$ 处取，左边较大的一个在 $x = a_n$ 处取，右边较大的一个在 $x = b_n$ 处取，即

$$a_n < c_n < b_n, \quad f(c_n) > f(a_n), \quad f(c_n) > f(b_n).$$

3）令 $y_1, y_2, \cdots, y_n$ 中最大的三个数为

$$\eta_1 = f(\xi_1) > \eta_2 = f(\xi_2) > \eta_3 = f(\xi_3)$$

由（2）可知 $\xi_1 = c_n$，ξ_2 是 a_n，b_n 之一，ξ_3 或是 a_n，b_n 之另一个，或在 a_n 之左 或在 b_n 之右．由这三个值作

$$\Delta\xi_3 = \xi_3 - \xi_2,$$

$$A\xi_3 = \frac{1}{2}(\xi_3 + \xi_2),$$

$$\Delta\eta_3 = \eta_3 - \eta_2,$$

$$\Delta\xi_2 = \xi_2 - \xi_1,$$

$$A\xi_2 = \frac{1}{2}(\xi_2 + \xi_1),$$

$$\Delta\eta_2 = \eta_2 - \eta_1.$$

再算出（参看 §1，(1)）

$$x_{n+1}^* = \frac{A\xi_3\Delta\xi_3\Delta\eta_2 - A\xi_2\Delta\xi_2\Delta\eta_3}{\Delta\xi_3\Delta\eta_2 - \Delta\xi_2\Delta\eta_3} \tag{1}$$

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

4）x_{n+1} 的定义如下

$$
x_{n+1}=\begin{cases}
x_{n+1}^{*}, & \text{若 } a_n < x_{n+1}^{*} < b_n,\ x_{n+1}^{*} \neq c_n. \\
x_{n+1}^{*}-\varepsilon, & \text{若 } x_{n+1}^{*}=c_n,\ c_n-a_n \geqslant b_n-c_n. \\
x_{n+1}^{*}+\varepsilon, & \text{若 } x_{n+1}^{*}=c_n,\ b_n-c_n > c_n-a_n. \\
\frac{1}{2}(a_n+c_n), & \text{若 } x_{n+1}^{*} \leqslant a_n. \\
\frac{1}{2}(c_n+b_n), & \text{若 } x_{n+1}^{*} \geqslant b_n.
\end{cases}
$$

试验出 $y_{n+1}=f(x_n+1)$ 的值，归纳法继续进行，这里 ε 是个适当小的数，但应使 $x_{n+1}^{*}+\varepsilon$ 或 $x_{n+1}^{*}-\varepsilon$ 与 x_{n+1}^{*} 的试验结果有所区别.

附记 1）由以上的理论研究可知，到充分多次后，实质上总是 $x_{n+1}=x_{n+1}^{*}$.

2）公式（1）的几何意义如下：令

$$
X=\frac{A\xi_3\Delta\xi_3\Delta\eta_2-A\xi_2\Delta\xi_2\Delta\eta_3}{\Delta\xi_3\Delta\eta_2-\Delta\xi_2\Delta\eta_3},
$$

即得

$$
\frac{X-A\xi_3}{\Delta\xi_2}\Delta\eta_2=y=\frac{X-A\xi_2}{\Delta\xi_3}\Delta\eta_3
$$

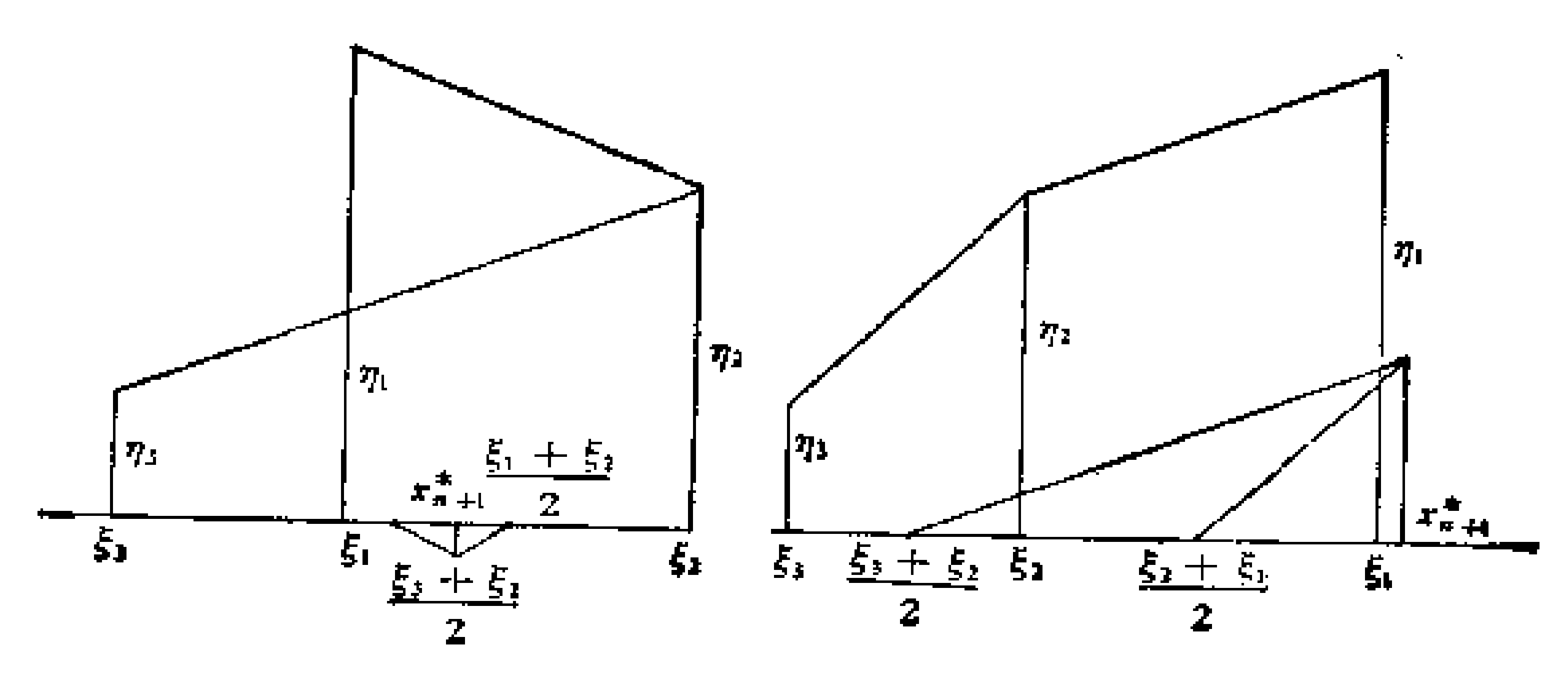


图 12

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

左边表示通过$\left(\frac{1}{2}(\xi_3+\xi_2), 0\right)$且平行于 (ξ_1, η_1), (ξ_2, η_2)的联线的直线，右边表示通过$\left(\frac{1}{2}(\xi_1+\xi_2), 0\right)$且平行于$(\xi_2, \eta_2)$, (ξ_3, η_3)的联线的直线，这两条直线的交点的横坐标就是x_{n+1}^*。

这就是从抛物线上三点找出顶点的方法。

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

第三章　分批试验及其他

述　　要

在生产和科学实验中，往往有这样的情况：做完一个试验要较长的时间才能看到结果．若等上次试验结果出来后，才进行下次试验，则要花费较多的时间．如果试验还要进行多次，那末消耗的时间将更多．为加速试验的进行，常常采用一批同时做几个试验的方法．

例如，一批做四个试验，共做两批．我们可以用如下的三种方法安排：

1)　○○＊○○＊○○＊○○＊○○

2)　○○○＊○＊○○○＊○＊○○○

3)　○○○○＊＊○○○○＊＊○○○○

在打＊的四点做第一批试验，然后在较好的一点附近四个打○的点处做第二批试验．我们看到第三种方法共区别 16 个点，可以区别的点数最多，但由于两对试验点的距离很近，试验的结果可能不易辨别好坏，甚至会出现这样的情况：由于观察的误差，好的一段反而被去掉了．第一种方法少区别两个点，但第一批试验的点分布均匀，可以避免上述情况，所以人们常常采用第一种方法．

对于每批做偶数个试验的安排是类似的．如果每批试验的个数是奇数，则安排便有所不同了，例如，每批做三个试验，做两批，也有三种方法：

1)　○○○＊＊○○○＊

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

2）　○○○＊＊＊○○○

3）　○○＊○＊○○＊○

它们都可以区别九个点，但每种方法都有其特点，与每批做四个试验的情况类似，第一种方法的第四点、第五点试验结果可能不易分辨，而第九点则没有充分发挥作用，如果这么靠近还可以分辨，那末还是采用第二种方法好，因为在第一批试验中，如果第五点的结果好，它便是最好点，省去一批试验．第三种方法是为了避免出现试验点太靠近而考虑的，不过，如果要考虑做三批以上的试验，则第一种方法又是常采用的．

从上述二例说明，在一批做好几个试验的时候，必须根据实际情况，选择适当的安排方法．

§1. 试验方案优劣的衡量标准

在衡量分批试验方案优劣时，往往"一般地"采用留下区间的长度或最好试验点距留下区间端点的最大距离作为误差（或精度），使之最小．在前面的假定下，我们在 $[0, 1]$ 区间中点附近相距极小的两点做试验：

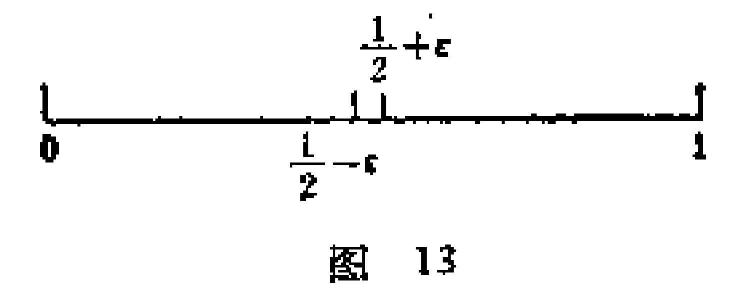


图　13

便可去掉区间的 $1/2-\varepsilon$．于是，也就不可避免地引入"在同一点作两个试验可以去掉区间一半"等这样纯数学的假设．

我们从实际应用出发，要求：最后一批试验点和上批留下的较好点等分上批试验后所留下的区间．这样使得全部试验结束后，所找到的最好试验点，距留下区间端点距离相等．在上述前提下，我们用最好试验点与留下区间端点的距离作

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

为误差，并使之最小。

§2. 一组特殊的方程组的解法

给出方程组

$$\left\{\begin{aligned} 1 = a_0 &= \lambda_1 a_1 + \mu_1 a_2, \\ a_1 &= \lambda_2 a_2 + \mu_2 a_3, \\ a_2 &= \lambda_3 a_3 + \mu_3 a_4, \\ &\cdots\cdots\cdots\cdots\cdots \\ a_{l-1} &= \lambda_l a_l + \mu_l a_{l+1}, \\ a_l &= a_{l+1}, \end{aligned}\right. \tag{1}$$

其中 $\lambda_1, \lambda_2, \cdots, \lambda_l; \mu_1, \mu_2, \cdots, \mu_l$ 为已知数，$a_1, a_2, \cdots, a_{l+1}$ 为待求数。

此方程组固然可以用最后一方程代入上一方程，解出 $a_{l-1} = (\lambda_l + \mu_l)a_{l+1}$，然后又代入前一方程，$\cdots$，最后解出 a_{l+1}。此方法较烦琐，我们可采用如下巧妙的方法：

令

$$\frac{a_m}{a_{m+1}} = \nu_m, \tag{2}$$

则 (1) 化为

$$\left\{\begin{aligned} \frac{1}{a_1} = \nu_0 &= \lambda_1 + \frac{\mu_1}{\nu_1} \\ \nu_1 &= \lambda_2 + \frac{\mu_2}{\nu_2} \\ \nu_2 &= \lambda_3 + \frac{\mu_3}{\nu_3} \\ &\cdots\cdots\cdots\cdots \\ \nu_{l-1} &= \lambda_l + \frac{\mu_l}{\nu_l} \\ \nu_l &= 1 \end{aligned}\right. \tag{3}$$

由 (3)，可将 $\nu_0, \nu_1, \cdots, \nu_l$ 表为连分数：

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

$$
\begin{cases}
\nu_l = 1 \\
\nu_{l-1} = \lambda_l + \dfrac{\mu_l}{\nu_l} = \lambda_l + \mu_l \\
\nu_{l-2} = \lambda_{l-1} + \dfrac{\mu_{l-1}}{\nu_{l-1}} = \lambda_{l-1} + \dfrac{\mu_{l-1}}{\lambda_l - \mu_l} \\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
\nu_2 = \lambda_3 + \dfrac{\mu_3}{\lambda_4} + \dfrac{\mu_4}{\lambda_5} + \cdots + \dfrac{\mu_{l-1}}{\lambda_l + \mu_l} \\
\nu_1 = \lambda_2 + \dfrac{\mu_2}{\lambda_3} + \dfrac{\mu_3}{\lambda_4} + \cdots + \dfrac{\mu_{l-1}}{\lambda_l + \mu_l} \\
\nu_0 = \lambda_1 + \dfrac{\mu_1}{\lambda_2} + \dfrac{\mu_2}{\lambda_3} + \cdots + \dfrac{\mu_{l-1}}{\lambda_l + \mu_l}
\end{cases}
\qquad (4)
$$

令

$$
F_0 = F_1 = 1, \quad F_{m+1} = \lambda_{l-m+1} F_m + \mu_{l-m+1} F_{m-1} \qquad (5)
$$
$$
(m = 1, 2, \cdots, l)
$$

则由 (4) 得

$$
\begin{cases}
\nu_l = \dfrac{F_1}{F_0} \\
\nu_{l-1} = \dfrac{F_2}{F_1} \\
\nu_{l-2} = \dfrac{F_3}{F_2} \\
\cdots\cdots\cdots \\
\nu_2 = \dfrac{F_{l-1}}{F_{l-2}} \\
\nu_1 = \dfrac{F_l}{F_{l-1}} \\
\nu_0 = \dfrac{F_{l+1}}{F_l}
\end{cases}
\qquad (6)
$$

由 (2) 和 (6),

$$
a_k = \frac{1}{\nu_0 \cdot \nu_1 \cdots \nu_{k-1}} = \frac{F_l}{F_{l+1}} \cdot \frac{F_{l-1}}{F_l} \cdots\cdots \frac{F_{l-k+1}}{F_{l-k+2}}
$$

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

$$= \frac{F_{l-k+1}}{F_{l+1}}$$

$(k = 1, 2, \cdots, l+1)$，即有

$$\begin{cases} a_0 = 1 = F_{l+1} \cdot a_l \\ a_1 = \dfrac{F_l}{F_{l+1}} = F_l \cdot a_l \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ a_k = \dfrac{F_{l-k+1}}{F_{l+1}} = F_{l-k+1} \cdot a_l \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ a_{l-1} = \dfrac{F_2}{F_{l+1}} = F_2 \cdot a_l \\ a_l = a_{l+1} = \dfrac{1}{F_{l+1}} \end{cases} \tag{7}$$

§3. 每批作奇数个试验，如何安排

试验范围 $a_0 = 1$，共做 l 批试验第一批 k_1 个试验，第二批 k_2 个试验，$\cdots$，第 l 批 k_l 个试验，其中 $k_1, k_2, \cdots, k_l$ 均为奇数. 试验安排方案如下：

把 a_0 分为长度为 $a_1, a_2 (a_1 \geqslant a_2)$ 互相间隔的 (k_1+1) 个小区间：

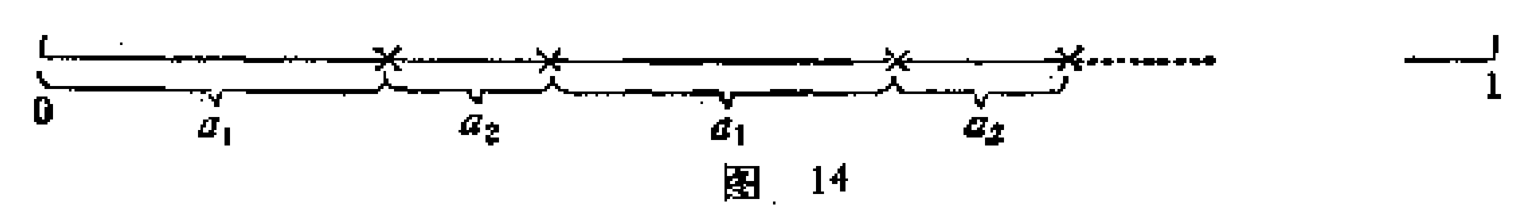


图 14

试验后留下包含取较大值试验点的区间：

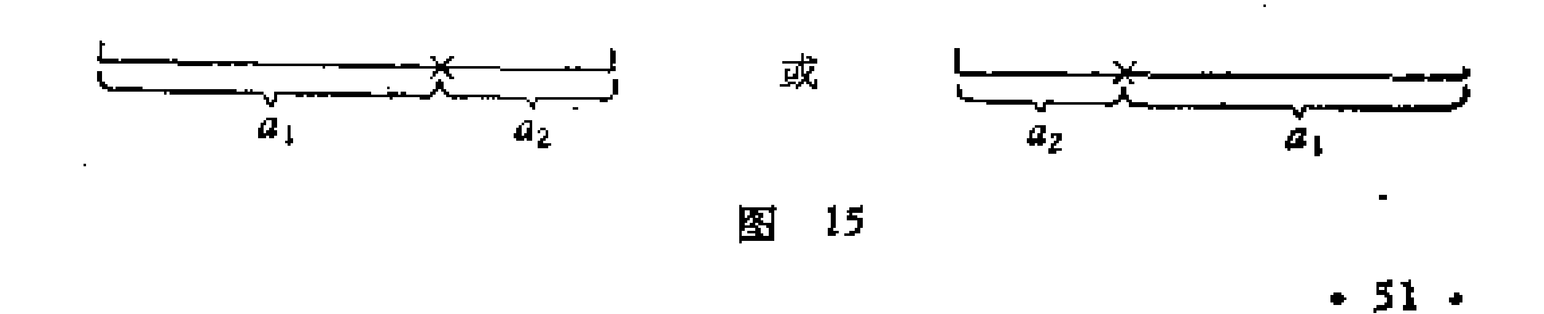


图 15

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

第二批 k_2 个试验全部安排在长度为 a_1 的区间里，并把 a_1 分成长度为 a_2，$a_3(a_2 \geqslant a_3)$ 互相间隔的 (k_2+1) 个小区间：

$$a_2a_3a_2a_3\cdots a_2a_3 * a_2 \quad \text{或} \quad a_2 * a_3a_2\cdots a_3a_2$$

试验后又留下包含取较大值试验点的区间：

$$a_2a_3 \quad \text{或} \quad a_3a_2$$

把下批试验全部安排在较长区间里，并把其区间分为长短相间的 (k_3+1) 个小区间，等等。根据 § 1 的要求，第 l 批的 k_l 个试验点要将 a_{l-1} 分为与 a_l 相等的 (k_l+1) 个小区间，最后留下

图 16

此处的 a_l 即为我们所定义的误差。

把上述的 $\{a_m\}$ 的关系写为数学表达式，就是 § 2 方程 (1) 的形式，其中

$$\begin{cases} \lambda_m = \dfrac{k_m+1}{2} \quad (m=1,2,\cdots,l), \\ \mu_m = \dfrac{k_m+1}{2} \quad (m=1,2,\cdots,l). \end{cases} \tag{1}$$

根据 § 2 中 (7) 式可以算出 a_m，而误差 $a_l = \dfrac{1}{F_{l+1}}$，F_{l+1} 由下面的递推公式确定：

$$\begin{cases} F_0 = F_1 = 1, \\ F_{m+1} = \dfrac{k_{l-m+1}+1}{2}(F_m + F_{m-1}) \quad (m=1,2,\cdots,l). \end{cases} \tag{2}$$

下面举几个特例：

例 1 只作一批 k 个试验。

由 (2) 式得

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

$$F_0 = F_1 = 1, \quad F_2 = k + 1.$$

再由§2中(7)式知

$$a_0 = 1, \quad a_1 = \frac{1}{F_2} = \frac{1}{k+1},$$

即 k 个试验点把区间 a_0 分为 $(k+1)$ 等分，试验误差

$$a_2 = \frac{1}{k+1}.$$

例 2 共作 l 批，每批作一个试验

由(2)式得

$$\begin{cases} F_0 = F_1 = 1, \\ F_{m+1} = F_m + F_{m-1} \quad (m = 1, 2, \cdots, l). \end{cases}$$

这就是通常的分数法. 试验误差为 $a_l = \dfrac{1}{F_{l+1}}$.

例 3 每批作 $(2t-1)$ 个试验，共作 k 批.

这时(2)式变为

$$\begin{cases} F_0 = F_1 = 1, \\ F_{m+1} = t(F_m + F_{m-1}) \quad (m = 1, 2, \cdots, k). \end{cases}$$

用第一章§2的方法，可得

$$F_{m+1} = \frac{1}{2}\left(1 + 3\sqrt{\frac{t}{t+4}}\right)\left[\frac{t + \sqrt{t(t+4)}}{2}\right]^m + \frac{1}{2}\left(1 - 3\sqrt{\frac{t}{t+4}}\right)\left[\frac{t - \sqrt{t(t+4)}}{2}\right]^m. \quad (3)$$

且有

$$a_1 = \frac{F_k}{F_{k+1}},$$

$$a_2 = \frac{F_{k-1}}{F_{k+1}},$$

$$\cdots\cdots\cdots$$

$$a_k = a_{k+1} = \frac{1}{F_{k+1}}.$$

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

试验误差为 $a_{k+1}=\dfrac{1}{F_{k+1}}$.

§4. 每批作偶数个试验，如何安排

设第一批 k_1 个试验，第二批 k_2 个试验，…，第 l 批 k_l 个试验，k_1，k_2，…，k_l 皆为偶数. 把试验区间 $a_0(=1)$ 分为长度是 a_1，$a_{l+1}(a_1\geqslant a_{l+1})$ 互相间隔的 (k_1+1) 个小区间：

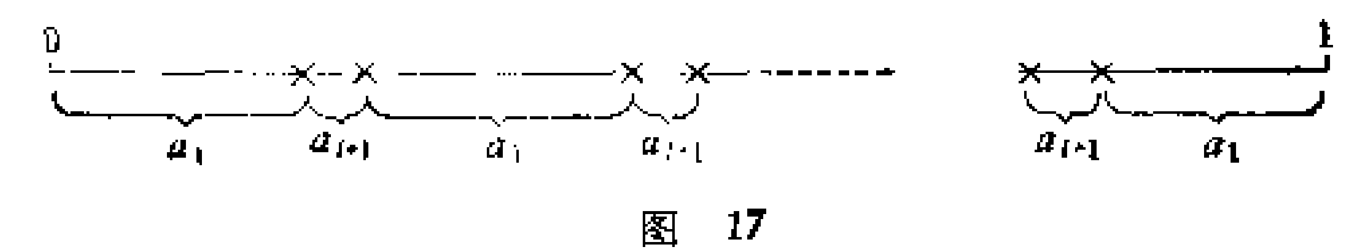


图 17

第二批试验全部安排在长为 a_1 的小区间，并把 a_1 分为长度 a_2，a_{l+1} $(a_2\geqslant a_{l+1})$ 互相间隔的 (k_2+1) 个小区间，等等. 根据 §1 的要求，第 l 批的 k_l 个试验要将区间 a_{l-1} 分为与 a_{l+1} 相等的 k_l+1 个小区间：

a_{l+1} a_{l+1}

图 18

此处 a_{l+1} 即为我们所定义的误差.

把上述的 $\{a_m\}$ 的关系写成数学表达式：

$$\begin{cases} 1=a_0=\left(\dfrac{k_1}{2}+1\right)a_1+\dfrac{k_1}{2}a_{l+1}, \\ \quad a_1=\left(\dfrac{k_2}{2}+1\right)a_2+\dfrac{k_2}{2}a_{l+1}, \\ \quad \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{l-1}=\left(\dfrac{k_l}{2}+1\right)a_l+\dfrac{k_l}{2}a_{l+1}, \\ \quad a_l=a_{l+1}. \end{cases} \tag{1}$$

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

或

$$a_{m-1}+a_{l+1}=\left(\frac{k_m}{2}+1\right)(a_m+a_{l+1}) \tag{2}$$

所以

$$1+a_{l+1}=a_0+a_{l+1}$$
$$=\left(\frac{k_1}{2}+1\right)\left(\frac{k_2}{2}+1\right)\cdots\left(\frac{k_l}{2}+1\right)(a_l+a_{l+1}).$$

$$a_{l+1}=\frac{1}{2\left(\frac{k_1}{2}+1\right)\left(\frac{k_2}{2}+1\right)\cdots\left(\frac{k_l}{2}+1\right)-1}. \tag{3}$$

再由

$$1+a_{l+1}=a_0+a_{l+1}=\left(\frac{k_1}{2}+1\right)\left(\frac{k_2}{2}+1\right)$$
$$\cdots\left(\frac{k_m}{2}+1\right)(a_m+a_{l+1}),$$

可以算出 $a_m(m=1,2,\cdots,l)$.

$$a_m=\left[2\left(\frac{k_{m+1}}{2}+1\right)\left(\frac{k_{m+2}}{2}+1\right)\cdots\left(\frac{k_{m+l}}{2}+1\right)-1\right]$$
$$\Big/\left[2\left(\frac{k_1}{2}+1\right)\left(\frac{k_2}{2}+1\right)\cdots\right.$$
$$\left.\left(\frac{k_l}{2}+1\right)-1\right]. \tag{4}$$

例　每批作 $2t$ 个试验，共做 l 批.

由(3)及(4)得

$$a_{l+1}=\frac{1}{2(t+1)^l-1},$$
$$a_m=\frac{2(t+1)^{l-m}-1}{2(t+1)^l-1}.$$

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

§5. 一般情形，如何安排.

和§3，§4方法一样．若第$(m-1)$批后留下区间为



图 19

$(a_{m-1}\geqslant a_p)$，则第m批试验点全部安排在较长区间a_{m-1}内．当k_m是奇数时，$a_p=a_m$；当k_m是偶数时，$a_p=a_{m+i}$．k_{m+i}是k_{m+1}，k_{m+2}，$\cdots$中出现的第一个奇数．

具体地说，设k_1，k_2，$\cdots$，k_l中的奇数依次为k_{i_1}，k_{i_2}，$\cdots$，k_{i_t}，则各区间长度的数学表达式为：

$$
\begin{cases}
1=a_0=\left[\dfrac{k_1}{2}+1\right]a_1+\left[\dfrac{k_1}{2}\right]a_{i_1} \\
\cdots\cdots\cdots\cdots\cdots\cdots \\
a_{i_1-2}=\left[\dfrac{k_{i_1-1}}{2}+1\right]a_{i_1-1}+\left[\dfrac{k_{i_1-1}}{2}\right]a_{i_1}, \\
a_{i_1-1}=\left[\dfrac{k_{i_1}}{2}+1\right]a_{i_1}+\left[\dfrac{k_{i_1}}{2}+1\right]a_{i_2}, \\
a_{i_1}=\left[\dfrac{k_{i_1+1}}{2}+1\right]a_{i_1+1}+\left[\dfrac{k_{i_1+1}}{2}\right]a_{i_2}, \\
\cdots\cdots\cdots\cdots\cdots\cdots \\
a_{i_2-2}=\left[\dfrac{k_{i_2-1}}{2}+1\right]a_{i_2-1}+\left[\dfrac{k_{i_2-1}}{2}\right]a_{i_3}, \\
a_{i_2-1}=\left[\dfrac{k_{i_2}}{2}+1\right]a_{i_2}+\left[\dfrac{k_{i_2}}{2}+1\right]a_{3}, \\
\cdots\cdots\cdots\cdots\cdots\cdots \\
a_l=a_{l+1}.
\end{cases}
\qquad (1)
$$

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

同§4一样，可以把偶数次的试验合并写成

$$
\begin{cases}
1 = a_0 = \left[\frac{k_1}{2}+1\right]\left[\frac{k_2}{2}+1\right]\cdots \\
\qquad \left[\frac{k_{i_1-1}}{2}+1\right](a_{i_1-1}+a_{i_1}) - a_{i_1}, \\
a_{i_1-1} = \left[\frac{k_{i_1}}{2}+1\right]a_{i_1} + \left[\frac{k_{i_1}}{2}+1\right]a_{i_2}, \\
a_{i_1} = \left[\frac{k_{i_1+1}}{2}+1\right]\left[\frac{k_{i_1+2}}{2}+1\right]\cdots \\
\qquad \left[\frac{k_{i_2-1}}{2}+1\right](a_{i_2-1}+a_{i_2}) - a_{i_2}, \\
a_{i_2-1} = \left[\frac{k_{i_2}}{2}+1\right]a_{i_2} + \left[\frac{k_{i_2}}{2}+1\right]a_{i_3}, \\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
a_{i_l} = \left[\frac{k_{i_l+1}}{2}+1\right]\left[\frac{k_{i_l+2}}{2}+1\right]\cdots \\
\qquad \left[\frac{k_l}{2}+1\right](a_l+a_{l+1}) - a_{l+1}, \\
a_l = a_{l+1}.
\end{cases} \tag{2}
$$

进一步可简化成(记 $i_{l+1}=l$)

$$
\begin{cases}
1 = a_0 = \lambda_1 a_{i_1} + \mu_1 a_{i_2}, \\
a_{i_1} = \lambda_2 a_{i_2} + \mu_2 a_{i_3}, \\
\cdots\cdots\cdots\cdots\cdots \\
a_{i_l} = \lambda_{l+1} a_l + \mu_{l+1} a_{l+1}, \\
a_l = a_{l+1}.
\end{cases} \tag{3}
$$

在这里，

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

$$
\begin{cases}
\lambda_m = \left[\dfrac{k_{i_{m-1}+1}}{2} + 1\right]\left[\dfrac{k_{i_{m-1}+2}}{2} + 1\right]\cdots \\
\qquad\quad \left[\dfrac{k_{i_m-1}}{2} + 1\right]\left[\dfrac{k_{i_m}}{2} + 2\right] - 1, \\
\mu_m = \left[\dfrac{k_{i_{m-1}+1}}{2} + 1\right]\cdots\left[\dfrac{k_{i_m-1}}{2} + 1\right]\left[\dfrac{k_{i_m}}{2} + 1\right], \\
\qquad (m = 1, 2, \cdots, t;\ i_0 = 0) \\
\lambda_{t+1} = \left[\dfrac{k_{i_t+1}}{2} + 1\right]\cdots\left[\dfrac{k_l}{2} + 1\right] - 1, \\
\mu_{t+1} = \left[\dfrac{k_{i_t+1}}{2} + 1\right]\cdots\left[\dfrac{k_l}{2} + 1\right].
\end{cases} \tag{4}
$$

用§2的方法，可以算出误差 a_l 以及 $a_{i_1}, a_{i_2}, \cdots, a_{i_t}$. 进而由(1)可以算出所有的 $\{a_m\}$.

令

$$
\begin{aligned}
&F_0 = F_1 = 1, \\
&F_{m+1} = \lambda_{t-m+1}F_m + \mu_{t-m+1}F_{m-1} \\
&\qquad (m = 1, 2, \cdots, t),
\end{aligned} \tag{5}
$$

则误差为

$$a_l = \frac{1}{F_{t+1}}.$$

读者可能会问，是否还有其他方法使得误差（或精密度）更小一些？这个问题的回答是否定的，对于单峰函数的全体而言，精度不能再好了。这里不再详细讨论，请读者自己证明。例如，可以参考第一章§5的方法，先建立一批不等式，然后用归纳法证明。

§6. 试验批数不定的情形

有时，虽然每批试验的次数一定，但是试验的批数并不能

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

预先给定，要求试验可以连续做下去，怎么办？

先讨论每批做偶数 $(2t)$ 个试验的情形.

由 § 4 的例，每批作 $2t$ 个试验，共做 k 批，第一个试验点在

$$a_1 = \frac{2(t+1)^{k-1}-1}{2(t+1)^k-1},$$

第二个试验点在

$$a_1 + a_{k+1} = \frac{2(t+1)^{k-1}}{2(t+1)^k-1}.$$

当 $k \to \infty$时，二试点位置都趋于 $\frac{1}{t+1}$. 由此可见，这个方法当 $k \to \infty$ 时是完全失败的. 也就是说，预先给定的批数 k 越大，两点越近，分辨起来越困难，正因为如此，我们建议用下面的方法来安排.

先把试验范围等分为 $(2t+1)$ 段，在 $2t$ 个分点上做第一批试验，比较结果，留下较好的点 * 及其左右一段，然后把这二段都等分为 $(t+1)$ 段，在分点处做第二批试验等等，下图表明了 $t=2$ 的情况:

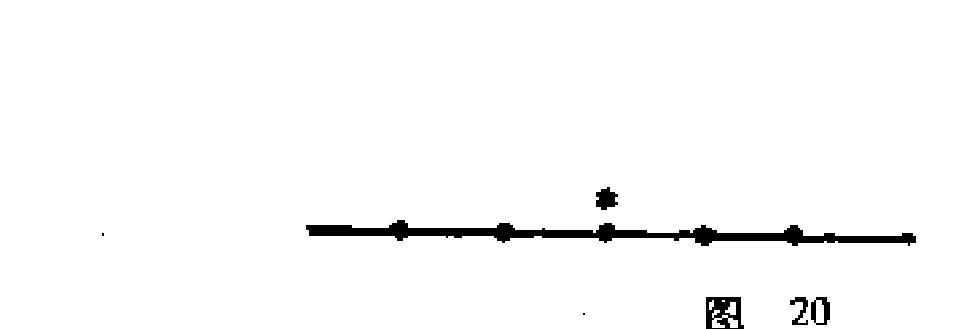


图 20

值得指出，根据实际情况，我们可以用混合的方法来安排，也就是先用 § 1 的方法，做到某一预定的程度后，还可细分时，再用上述方法.

再讨论每批做奇数 $(2t-1)$ 个试验的情形.

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

由§3的例3，每批作 $(2t-1)$ 个试验，共做 k 批，第一个试验点在

$$a_1=\frac{F_k}{F_{k+1}}=\left\{\frac{1}{2}\left(1+3\sqrt{\frac{t}{t+4}}\right)\left[\frac{t+\sqrt{t(t+4)}}{2}\right]^{k-1}\right.$$

$$\left.+\frac{1}{2}\left(1-3\sqrt{\frac{t}{t+4}}\right)\left[\frac{t-\sqrt{t(t+4)}}{2}\right]^{k-1}\right\}\Big/$$

$$\left\{\frac{1}{2}\left(1+3\sqrt{\frac{t}{t+4}}\right)\left[\frac{t+\sqrt{t(t+4)}}{2}\right]^{k}\right.$$

$$\left.+\frac{1}{2}\left(1-3\sqrt{\frac{t}{t+4}}\right)\left[\frac{t-\sqrt{t(t+4)}}{2}\right]^{k}\right\}.$$

当 $k\to\infty$ 时，它趋于

$$\theta=\frac{2}{t+\sqrt{t(t+4)}}=\frac{1}{2}\left(-1+\sqrt{\frac{t+4}{t}}\right).$$

同样可以得出，第二个试验点与第一个试验点的距离

$$a_2=\frac{F_{k-1}}{F_{k+1}}.$$

当 $k\to\infty$ 时，它趋于

$$\theta^2=\left[\frac{2}{t+\sqrt{t(t+4)}}\right]^2=\frac{1}{2}\left(\frac{t+2}{t}-\sqrt{\frac{t+4}{t}}\right).$$

容易看出

$$\theta+\theta^2=\frac{1}{t}.$$

这样，对每批做 $(2t-1)$ 个试验，我们可以按上式取 θ，即

$$\theta=\frac{1}{2}\left(-1+\sqrt{\frac{t+4}{t}}\right),$$

先把试验范围分成 $2t$ 段，其长度比是

$$\theta:\theta^2:\theta:\theta^2:\cdots:\theta:\theta^2.$$

在分点做第一批试验后，余下是

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

$$\theta:\theta^2 \quad 或 \quad \theta^2:\theta.$$

由于对称关系,只讨论前者.

因为

$$t(\theta^2+\theta^3)=t\theta(\theta+\theta^2)=\theta,$$

故把长为 θ 的一段分为 $2t$ 段,其长度比是

$$\theta^2:\theta^3:\theta^2:\theta^3:\cdots:\theta^2:\theta^3.$$

再在分点处做第二批试验,余下是

$$\theta^2:\theta^3 \quad 或 \quad \theta^3:\theta^2.$$

因为

$$t(\theta^3+\theta^4)=t(\theta+\theta^2)\theta^2=\theta^2,$$

又可把长为 θ^2 的一段分为 $2t$ 段,再做第三批试验,等等. 如此做下去,直到合乎要求为止.

这个方法是黄金分割法的推广,$t=1$ 时,就是黄金割法(一次做一个试验).

最后,看看这个方法的精度如何?

记 $l_k^{(i)}$ 为每批做 i 个试验,共做 K 批后余下的区间长度.对于 $i=2t$ 的情形,有

$$l_1^{(2t)}=\frac{2}{2t+1},$$

$$l_2^{(2t)}=\frac{2}{2t+1}\cdot\frac{1}{t+1},$$

$$\cdots\cdots\cdots\cdots\cdots\cdots$$

$$l_k^{(2t)}=\frac{2}{2t+1}\cdot\left(\frac{1}{t+1}\right)^{k-1}.$$

对于 $i=2t-1$ 的情形,有

$$l_1^{(2t-1)}=\theta+\theta^2=\frac{1}{t},$$

$$l_2^{(2t-1)}=\theta^2+\theta^3=\frac{\theta}{t},$$

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

$$l_3^{(2t-1)} = \theta^3 + \theta^4 = \frac{\theta^2}{t},$$

$$\cdots\cdots\cdots\cdots\cdots\cdots$$

$$l_k^{(2t-1)} = \theta^k + \theta^{k+1} = \frac{\theta^{k-1}}{t}.$$

§ 7. 是否最好的安排

上面几节讲的分批试验法是黄金分割法和分数法的自然推广，因而第一章里的许多概念和方法都可以平行移过来用.

现在我们对 § 6 的方法作进一步的讨论.

仍然先讨论每批做 $2t$ 个试验的情况.

我们的方法着眼于：做一批试验以后，进行比较，只留下一个区间，其中包含已经做过的、结果最好的那个试验点，也即如下图的形式. 我们将它放大适当倍数，使留下的区间为 $(0, 1)$，其中 θ 处是做过试验的点. 现在的问题是：再同时做 $2t$ 个试验，这 $2t$ 个点布置在何处，使下一次留下的区间尽可能短？



图 21

连同上次留下的已试验点，共有 $(2t+1)$ 个试验点，设安排在 $\theta_1, \theta_2, \cdots, \theta_{2t+1}$ 处，如果最好的在 θ_p，则下次试验将在

$$\theta_{p-1} < \theta < \theta_{p+1}$$

中做，因此下次试验的区间长度应当是

$$T = \max_{1 \leqslant p \leqslant 2t+1} (\theta_{p+1} - \theta_{p-1}), \quad \theta_0 = 0, \quad \theta_{2t+2} = 1.$$

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

由于

$$T=\max_{1\leqslant p\leqslant 2t+1}(\theta_{p+1}-\theta_{p-1})\geqslant\max_{1\leqslant r\leqslant t+1}(\theta_{2r}-\theta_{2r-2})$$

$$\geqslant\frac{1}{t+1}\sum_{r=1}^{t+1}(\theta_{2r}-\theta_{2r-2})=\frac{1}{t+1}(\theta_{2t+2}-\theta_0)$$

$$=\frac{1}{t+1},$$

也就是最短的因子是 $\frac{1}{t+1}$，比这个数更小的就无法保证，显然，均分法即

$$\theta_{p+1}-\theta_p=\frac{1}{2(t+1)}.$$

就可达到这一目的，这就是 § 2 所说的方法.

当然，如果第一批试验不是按等分取点，而是更细致些，安排为

图 22

以后再按等分取点. 容易证明这是最好的安排，其中

$$b=\frac{1}{2(t+1)^2-1},\quad a=\frac{2t+1}{2(t+1)^2-1},$$

满足

$$(a+b)t+a=\frac{2t+2}{2(t+1)^2-1}\cdot t$$

$$+\frac{2t+1}{2(t+1)^2-1}=1.$$

以后余下

$$a:b\quad 或\quad b:a$$

把长为 a 的一段等分为 $(2t+1)$ 分，即每段长为 b，在分点上做试验，以后均按 § 2 的方法. 这样第一批试验后余下区

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

间长为

$$l_1 = a + b = \frac{2t+2}{2(t+1)^2 - 1} < \frac{2}{2t+1}$$

以后缩短的因子都是 $\frac{1}{t+1}$.

为了取点简便，而精度又不相差太大，我们建议采用 § 6 介绍的方法.

再讨论每批做 $(2t-1)$ 个试验的情形。

这时，可以证明 § 6 给出的方法是最好的. 为此，我们只需证明下面两个引理，然后按照第一章 § 5 的思考方法去论证.

记 $\Delta_k^{(2t-1)}$ 为 k 批试验后的精度，我们有

引理 1 当 $k \geqslant 1$ 时，恒有

$$\Delta_{k+1}^{(2t-1)} \leqslant \Delta_k^{(2t-1)} \leqslant t(\Delta_{k+1}^{2t-1} + \Delta_{k+2}^{2t-1}). \tag{1}$$

证 左边不等式是显然的，故只证右方的不等式.

记 (x_0, x_{2t+1}) 是第 k 批试验余下区间之最大者，即

$$x_{2t+1} - x_0 = \Delta_k^{(2t-1)}. \tag{2}$$

第 $(k+1)$ 批试验 $(2t-1)$ 个，连上次留下的已试验点，共有 $2t$ 个试验点，记为 $x_1, x_2, \cdots, x_{2t}$,

且

$$x_0 < x_1 < x_2 < \cdots < x_{2t} < x_{2t+1}.$$

我们有

$$x_{p+1} - x_{p-1} \leqslant \Delta_{k+1}^{(2t-1)} \quad (p = 1, 2, \cdots, 2t).$$

由 (2)

$$\Delta_k^{(2t-1)} = x_{2t+1} - x_0 = \sum_{r=0}^{t-1} (x_{2r+2} - x_{2r}) + (x_{2t+1} - x_{2t})$$

$$\leqslant t\Delta_{k+1}^{(2t-1)} + (x_{2t+1} - x_{2t}). \tag{3}$$

在区间 (x_{2t}, x_{2t+1}) 中，第 $(k+2)$ 批试验点最多只能安排 $(2t-1)$ 个，

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

$$x_{2k+1}-x_{2t}\leqslant t\Delta_{k+2}^{(2t-1)}. \tag{4}$$

由 (3) 和 (4) 便得 (1) 式.

引理 2 当 $k\geqslant 1$, $n\geqslant 1$ 时,有

$$\Delta_n^{(2t-1)}\leqslant t\left(\frac{F_k^{(t)}}{t}\Delta_{n+k}^{(2t-1)}+F_{k-1}^{(t)}\Delta_{n+k+1}^{(2t-1)}\right). \tag{5}$$

其中 $\{F_k^{(t)}\}$ 由递归公式

$$F_0^{(t)}=1,\quad F_1^{(t)}=t,\quad F_k^{(t)}=t(F_{k-1}^{(t)}+F_{k-2}^{(t)}) \tag{6}$$

给出.

由引理 1, 用归纳法易证.

对于 § 6 给出的安排方法,记 $l_k^{(2t-1)}$ 为第 k 批试验后留下的区间长度,容易验证

$$l_k^{(2t-1)}=t(l_{k+1}^{2t-1}+l_{k+2}^{2t-1}),\quad k\geqslant 1. \tag{7}$$

$$l_n^{(2t-1)}=t\left(\frac{F_k^{(t)}}{t}l_{n+k}^{(2t-1)}+F_{k-1}^{(t)}l_{n+k+1}^{(2t-1)}\right)\quad (k\geqslant 1, n\geqslant 1). \tag{8}$$

根据 (1), (5), (7), (8), 用第一章 § 5 的方法易证.

定理 对任一每批做 $(2t-1)$ 个试验的安排方法 $\mathscr{M}$, 记 Δ_k^{2t-1} 为第 k 批试验后的精度, 若 $\mathscr{M}$ 不同于 § 2 的方法,则一定有 $\mathscr{N}$, 使当 $K\geqslant\mathscr{N}$ 时,

$$\Delta_k^{(2t-1)}>l_k^{(2t-1)}=\frac{\theta^{k-1}}{t}.$$

即: § 2 的方法是最优的安排方法.

§ 8. 依某种要求进行试验

每批同时做几个试验虽然有加快速度的好处, 但是也有增加试验费用的坏处, 并且每批试验个数增多, 原材料消耗比按正比增加得还要快. 例如, 用分数法 89/144 做十次试验的效果和一批同时做 143 个试验的效果相等, 时间缩短为 1/10, 但试验的代价却是 14.3 倍,因此,必须权衡利弊来决定

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

采取何种方案. 此外,一批做多个试验时,试验范围缩小得很快,例如,每批做7个试验,共做六批,精度为2/19776,即为万分之一. 如上所述,两个试验点愈靠近,则通过试验来分辨它们也就愈困难,因此考虑批数做得很多,意义不大. 遇到这种情况,首先看看试验区间要求分得多么细,再决定一批做几个试验,共做几批,来安排试验点等等.

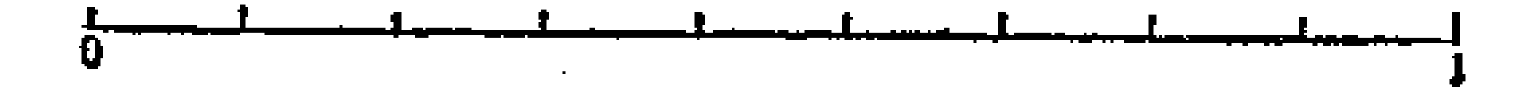
例如,有八套仪器,可以同时做试验,但在试验点比较接近时,要求重复试验来避免误差,简单地说,我们希望这样安排: 进行四批试验第一批八个试验点在八个不同的地方做,第二批四个试验点在四个不同的地方各做两个试验,第三批两个试验点在两个不同的地方各做四个试验,第四批在一点做六个试验,并在留下的较好点补做二个试验加以比较.也就是说每批试验次数依次为(8,4,2,1).

如果按§5的理论上推导,容易写出

$$\begin{cases} 1=a_0=5a_1+4a_4 \\ a_1=3a_2+2a_4 \\ a_2=2a_3+a_4 \\ a_3=a_4+a_5 \\ a_4=a_5 \end{cases}$$

并算出误差$a_5=1/89$. 但是,这样就要求第一批试验点两两相距1/89,与我们的前提"试验点比较接近时,要求重复试验来避免误差"不相符合了. 我们可以用下面的安排方法:

1) 先分试验范围为九等分,在5点上做八个试验(试验点相距1/9)

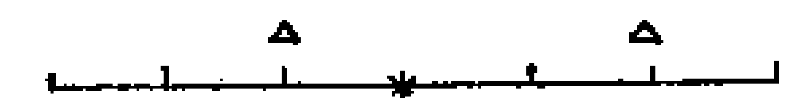


2) 对余下的区间仍按等分(试验点相距1/27)

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

每个打 * 处各做二个试验．

3）对余下的区间分成六等分，在第 2，5 分点各做四个试验：



4）最后在余下的一点做试验．

这样安排，误差是 $1/9 \times 1/3 \times 1/3 = 1/81$，与前面安排的误差 1/89 相差并不大，但却能有效地避免了两点相距很近不易分辨试验结果的困难．

§9. 重复性试验的分辨问题

我们在上一节说过，有时需要用重复试验来避免误差，由于在一个点 $x = a$，通过试验所得出的 $f(x)$ 的数据，并不一定是 $f(a)$，而是可能有误差的，例如测得的 $y = f(a) + \xi$，ξ 就是误差，为避免误差的严重影响，我们往往在同一条件 $x = a$ 的情况下，多做些试验，得出以下的数据

$$y_1, y_2, \cdots, y_t,$$

在另一点 $x = b$，也做了些试验，各得出

$$y'_1, y'_2, \cdots, y'_t.$$

如果 y_i 都比 y'_j 大，则两个试验是可以分辨的，也就是可以指望 $f(a) > f(b)$．但是，有些 y_i 比有些 y'_j 大，而有些 y_i 比有些 y'_j 小，怎么办？

第一个建议是用平均数法，如前，在 $x = a$ 时所测出的数据的平均数

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

$$\bar{y} = \frac{1}{s}(y_1 + y_2 + y_3 + \cdots + y_s)$$

比

$$\bar{y}' = \frac{1}{t}(y'_1 + y'_2 + \cdots + y'_t)$$

大，就认为 $f(a)$ 比 $f(b)$ 大，

第二个建议，看是否是

$$\begin{aligned} &\bar{y} > \text{每一个 } y'_j \quad (j = 1, 2, \cdots, t), \\ &\bar{y}' < \text{每一个 } y_i \quad (i = 1, 2, \cdots, s). \end{aligned} \tag{1}$$

如果是这样，我们认为 $f(a) > f(b)$ 的把握更大了。

第三个建议：考虑不适合不等式（1）的个数如果 y'_j，y_i 中有 p 个不适合（1），则我们可以用

$$1 - \frac{p}{s + t}$$

来检定 $f(a) > f(b)$。

第四个建议：(y_i, y'_j) 共有 st 对，其中适合于

$$y_i > y'_j$$

的有 q 对，则 $f(a) > f(b)$ 的把握是

$$\frac{q}{st}.$$

例如在两点各做了四个试验，得出数据

$$\begin{aligned} &y_1 = 1.19, \quad y_2 = 1.21, \quad y_3 = 1.21, \quad y_4 = 1.22, \\ &y'_1 = 1.10, \quad y'_2 = 1.15, \quad y'_3 = 1.18, \quad y'_4 = 1.20, \end{aligned}$$

我们只要依大小排队

$$y'_1 < y'_2 < y'_3 < y_1 < y'_4 < y_2 = y_3 < y_4,$$

y_i 与 y'_j 共有 $4 \times 4 = 16$ 种关系，15 个都适合 $y'_j < y_i$，只有一个 $y_1 < y'_4$，因此 $f(a) > f(b)$ 的可靠性是 $15/16 = 0.9375$。

上述这些处理方法，是比较简便而又避免了不必要的繁琐的处理方法，是比较容易掌握的。但在同一点反复做试验

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

得到数据很多的时候，我们不排斥运用"正态分布"（如果数据大致服从此规律时）中的标准离差等方法来判断最优点出现的可能性的范围.

还有一点值得注意，$f(a)$ 接近于 $f(b)$. 难以分辨时，即使就是 $f(a)=f(b)$ 时（或差不多相等），$f(x)$ 的最大值在 (a, b) 之间的可能性也是很大的，换言之，a 之左，b 之右都可以舍弃. 因而建议，暂且不管 (a, b) 之外，只对 (a, b) 开始用黄金分割法或其他方法，这样可以省做不少试验.

§10. 非单峰的情况如何办

有人会问，前面所讲的只适合于"单峰"的情况，多峰（即有几个点，其附近都比它差）的情况怎么办？我们建议：

1. 先不管它是单峰还是多峰，就按单峰的方法去做. 找到一个"峰"后，如果符合要求，就先照它生产，然后有时间继续再寻找其它可能的更高的"峰"（即分区寻找）.

2. 先做一批分布得比较均匀疏属的试验，看其是否有"多峰"的现象出现，如果有，则分区寻找. 这时，第一批试验的点最好依以下的比例划分

$$\alpha:\beta=0.618:0.382$$

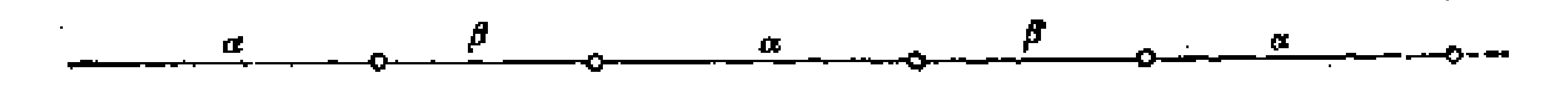


图 23

例如，三个分点，可以取之如

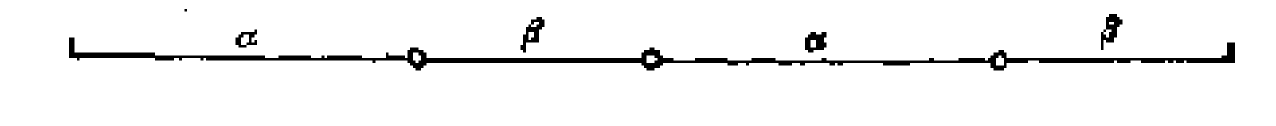


图 24

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

使留下的成为



图 25

的形式，接下去便可用 0.618 法了。

最后要说明一下，有些问题，并不是寻找单峰函数的最大值，而是如对分法那样，寻找一个合适的点。在区间 (a, b) 中有一个单调增加函数 $f(x)$，还有一个数 c，

$$f(a) < c < f(b).$$

要寻找 x_0 使

$$f(x_0) = c.$$

我们知道，对分法是在区间的中点做试验，若 $f\left(\frac{a+b}{2}\right) > c$，则舍弃 $\left(\frac{a+b}{2}, b\right)$；若 $f\left(\frac{a+b}{2}\right) < c$，则舍弃 $\left(a, \frac{a+b}{2}\right)$，再在余下区间的中点做试验，如此继续做几次试验后，余下区间的长度是

$$\frac{1}{2^n}(b-a).$$

容易证明，对于全体单峰函数来说，这个办法是最好的。如果条件允许每批做几个试验，那又如何安排呢？很简单，不管 m 是奇数还是偶数，这时都是用均分法最好，例如每批做三个实验，就把 (a, b) 四等分，

a θ_3 θ_2 θ_1 b

图 26

在三个分点 θ_1, θ_2, θ_3 处做，然后再考察试验的结果，例如

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

$f(\theta_2) < c$，$f(\theta_3) > c$，则舍弃 (a, θ_2) 和 (θ_3, b)，再四等分 (θ_2, θ_3)，在分点做第二批试验，如此继续做 n 批试验，余下的含有 x_0 的区间长度是

$$\frac{1}{4^n}(b-a).$$

也容易证明这个安排方法是最好的。

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

第二部分　多因素优选法

这一部分考虑的是多因素优选法，我们将分章研究这些方法在多因素情况下的推广． 和前面一样，每章都有一段方法的述要，在明确了方法的使用以后，再逐步深入研究方法的收敛性和一些有关的数学问题，最后一章是讨论优选法与计算数学的关系．所介绍的优选法，实质上就是求最大值(或最小值)的快速计算法．不仅如此，很多计算问题都可以化为极值问题来解决，而极值问题离散化后，就与优选法有关．因而优选法在计算数学上也是很重要的．

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

第一章 双因素优选法

述 要

绝大部分的双因素优选法都可以推广到两个以上因素的情况，只是所需的工作量，将随着因素个数的增加而迅速增长．下面我们先讲方法：

1）网格法　把因素 x 和 y 的优选范围(区域)

$$a < x < b, \quad c < y < d$$

各分为 l 和 m 个等分，于是总共有 lm 个格子，在每个格子中做一个试验，共做 lm 个试验后，哪个好用哪个．这个办法很简单，但它的工作量比起一个因素来说，是大大增加了．例如 $l=10$，$m=10$，就要做 100 个试验．

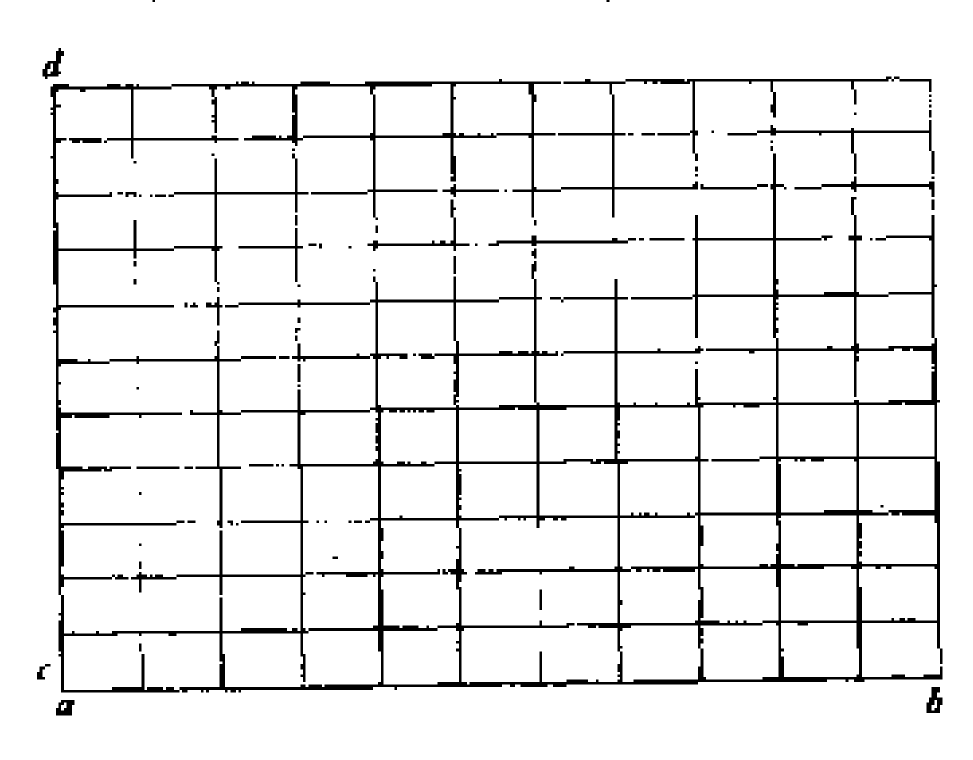


图 27

假定处理的是单峰问题，也就是把 x，y 平面作为水平面，试验出来的数值 z 看成为这一点的高度．这样的图形就

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

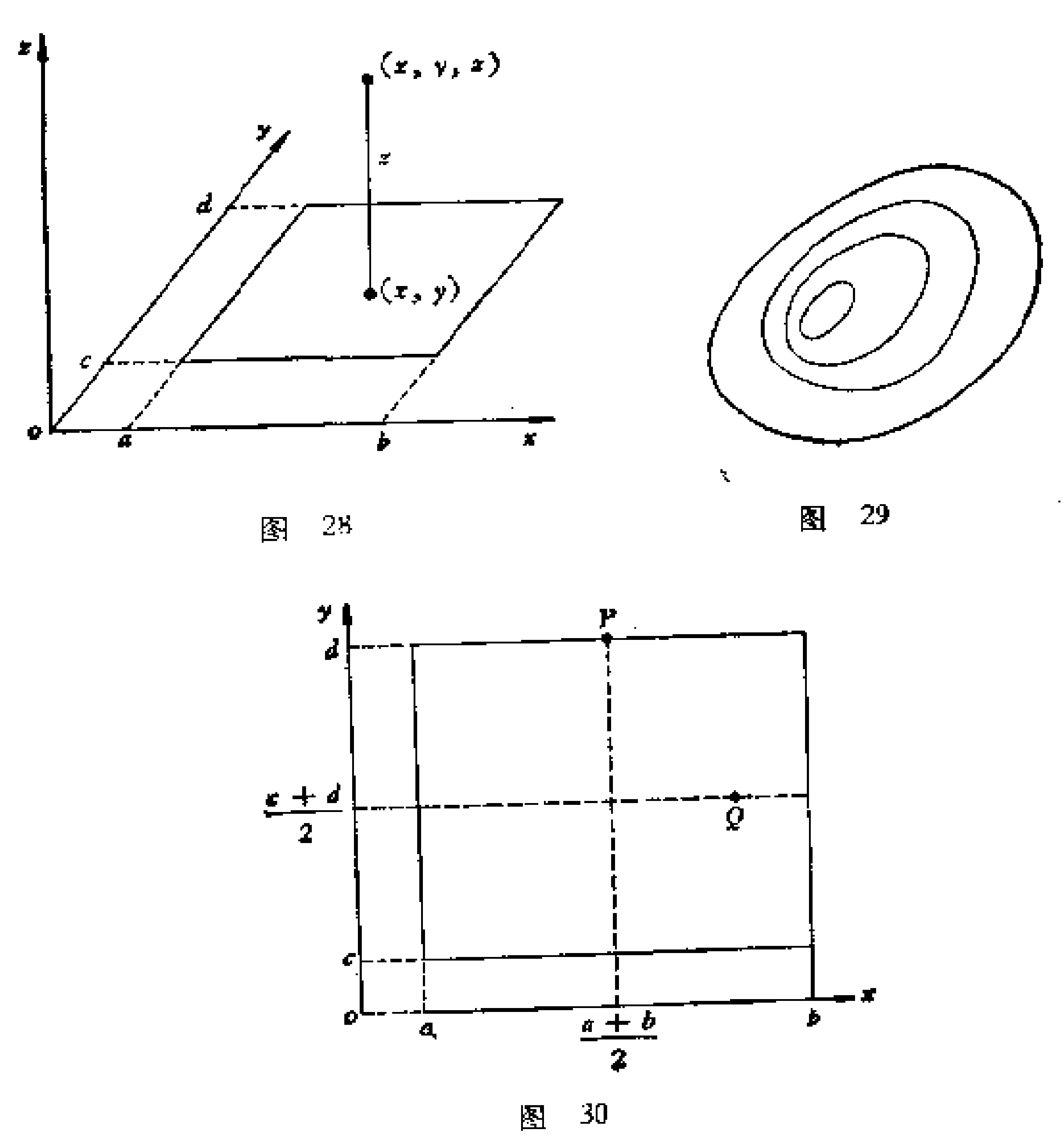


图 28　　图 29

图 30

像一座高山，我们假定这一片地区只有一个制高点，在平面上画出等高线后，这些等高线形成一圈圈的互不相交的封闭曲线．我们还假定，每一圈都是凸的，就和树木的年轮一样．

2）对开法　裁一块方格纸代表优选范围

$$a<x<b,\quad c<y<d.$$

左右对折起来，在中线

$$x=\frac{1}{2}(a+b)$$

上，用单因素法找最大值，设在 P 点取这最大值．再上下对

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

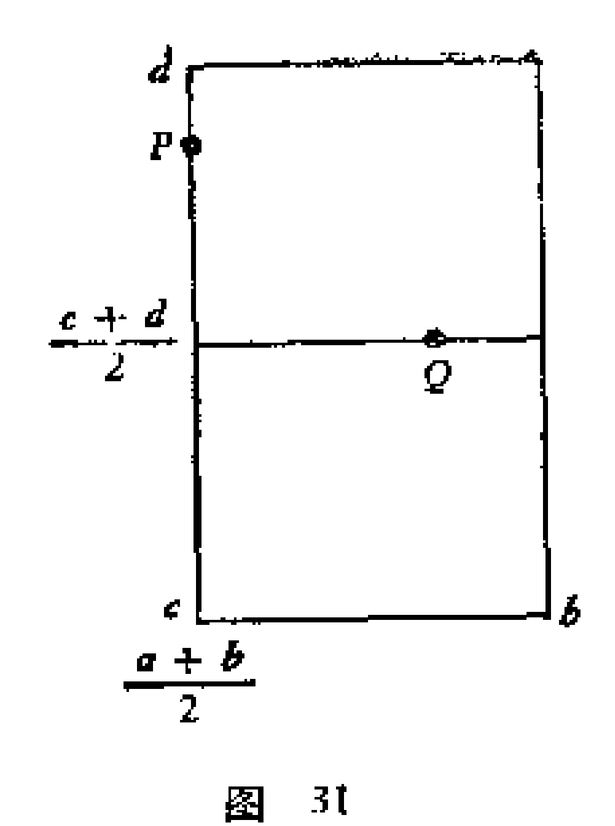


图 31

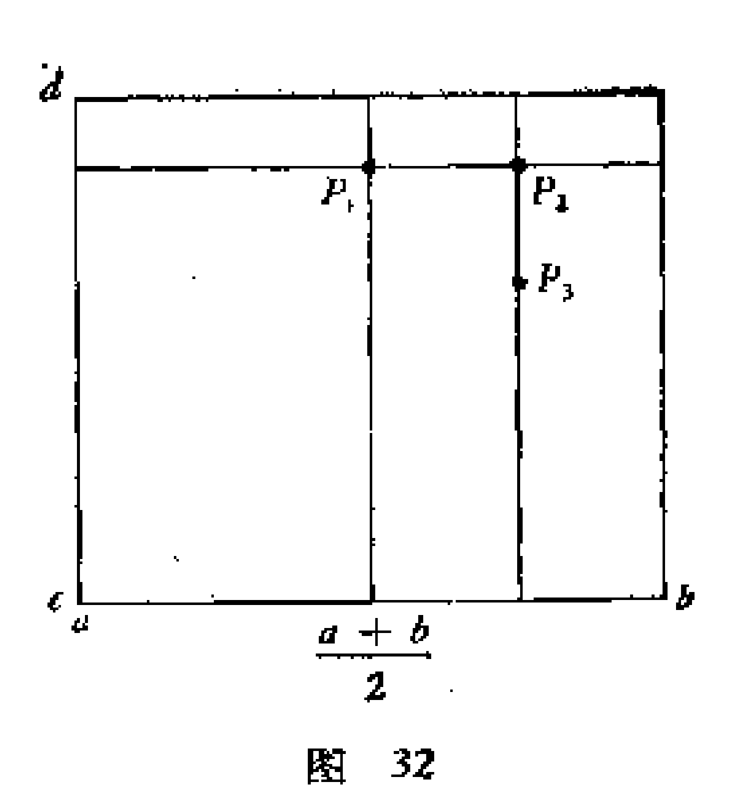


图 32

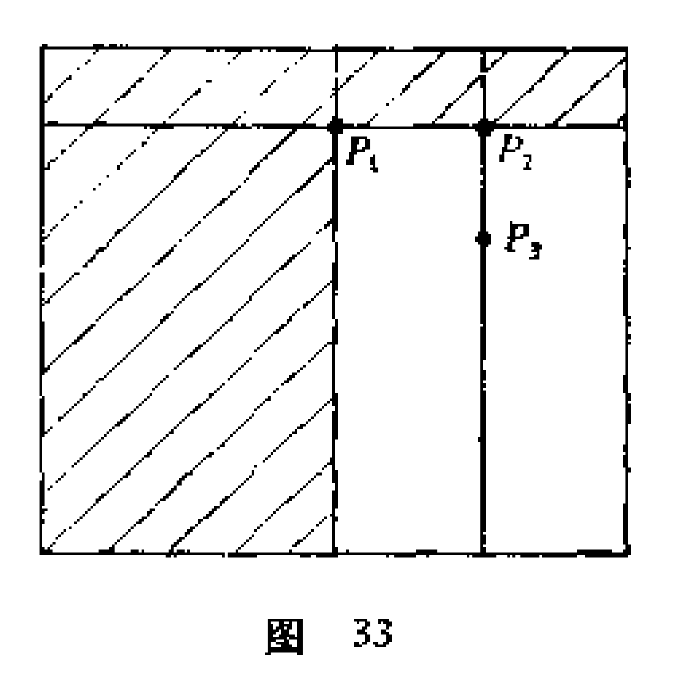


图 33

折，在中线

$$y=\frac{1}{2}(c+d)$$

上，用单因素法找最大值，假定在 Q 点取这值.

比较 P 点和 Q 点的结果，如果 Q 大，裁掉纸的左半张(不然裁掉下半张). 再用同样的方法来处理余下的半张纸，不断地去其一半，逐步地得出所需要的结果.

3) 旋升法　先在一条中线(例如左右对折线)上求最大值，假定在 P_1 点取这最大值. 通过 P_1 作水平线，在这水平线

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

上找最大值．假定在 P_2 处取，又在通过 P_2 的垂线上找最大值，而在 P_3 取这最大值，等等．继续做下去，有荫影的部分以后不要再考虑了．

4）平行线法　两个因素中，一个例如 x 易于调整，另一个例如 y 不易调整，则建议用“平行线法”，先把 y 固定在范围 (c, d) 的 0.618 处，即取

$$y = c + (d - c)0.618,$$

用单因素法找最大值，假定在 P 点取这值，再把 y 固定在范围 (c, d) 的 0.382 处，即取

$$y = c + (d - c)0.382,$$

用单因素法找最大值，假定在 Q 点取这值，比较 P，Q 的结果，如果 P 大，则去掉下面部分，即去掉

$$y \leqslant c + (d - c)0.382$$

的部分（否则去掉上面的部分），再用同样方法处理余下的部分，如此继续．这个方法也适用于分批试验的情况．

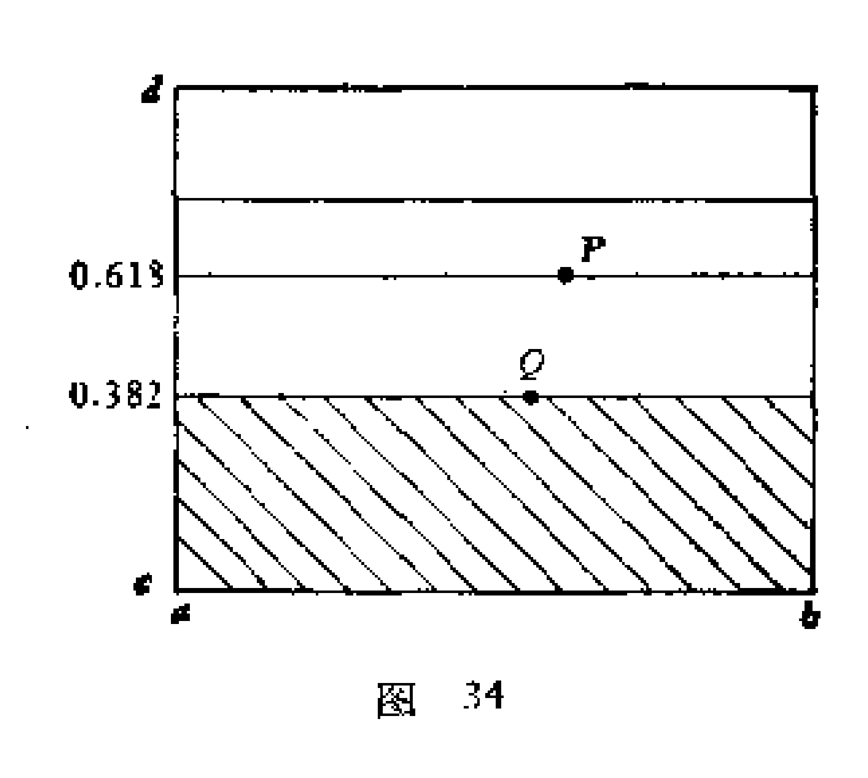


图　34

5）按格上升法　把所考虑的区域打上格子，与 2)，3)，4) 类似，但用分数法来代替黄金分割法．

6）翻筋斗法　从一个等边三角形 ABC 出发（如图 36.），在顶点各做一个试验，如果在 C 点所做的试验最好，则作 C 点的对顶同样大的等边三角形 CDE．在 D，E 处做试验，如果 D 点最好，则再作 D 点的对顶同样大的等边三角形，…，一直做下去．如果在 F，G 处做试验，都没有 D 点好，则取 FD 及 GD 的中点 F'，G' 做试验，也可以取 CD 及 ED 的中点做试

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

验，再用以上的方法．如果在D的两边一分再分都没有D点好，一般说来，D就是最好点了．

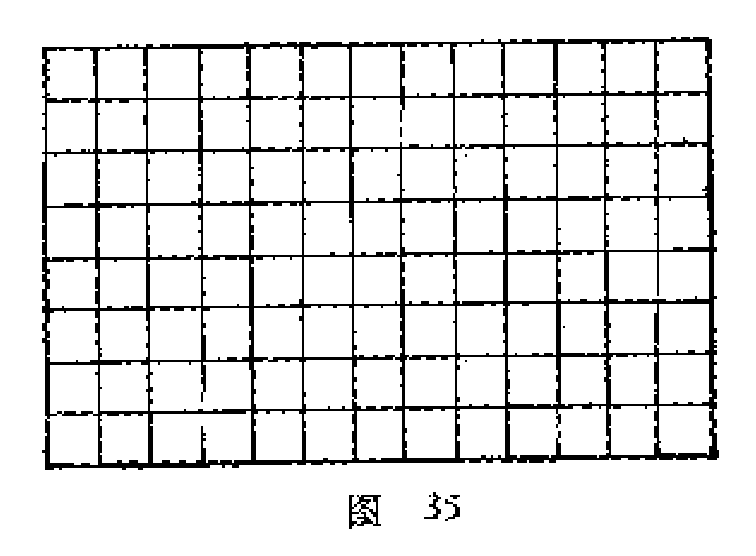


图 35

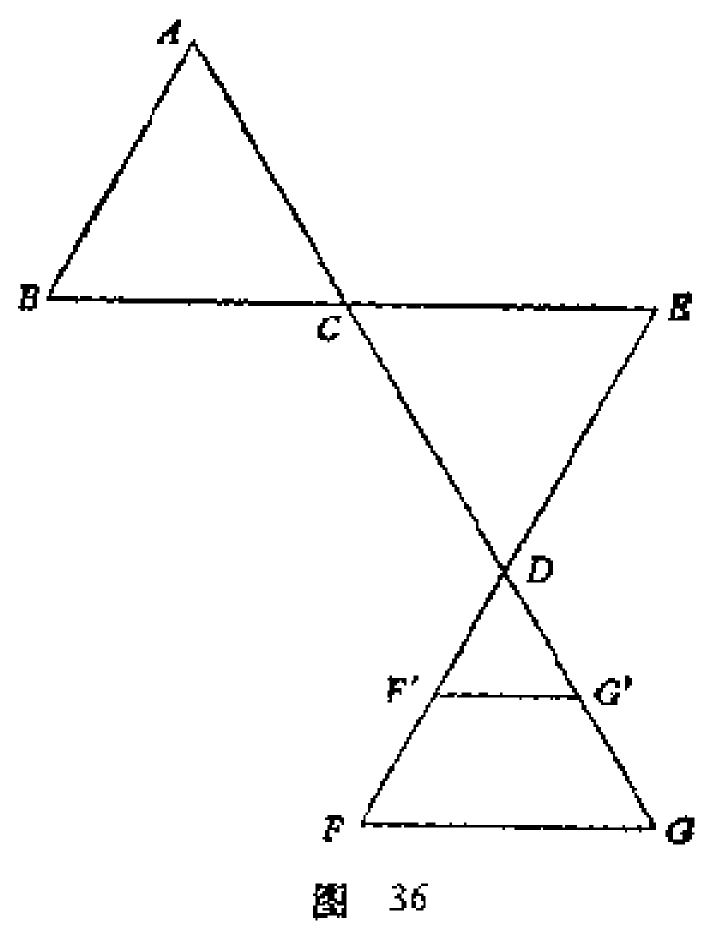


图 36

§1. 对 开 法

实际上，我们的优选问题就是在区域

$$a < x < b, \quad c < y < d$$

内求 $z = f(x, y)$ 的最大值．所谓等高线，就是固定 z，由方程 $z = f(x, y)$ 在 (x, y) 平面上所定义的曲线．我们假定等

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

高线有连续切线。

在 (x, y) 平面上，一条直线上取最大值 e 的点，一定是等高线 $f(x, y)=e$ 与这直线的切点，因为这直线不可能与它再交于另一点。由于等高线是凸的，所以，使 $z>e$ 的点一定分布在这直线的一侧，现在用这个原理来观察对开法。

(1) 这方法一定可以逐步通过最大值(或简称收敛法)。

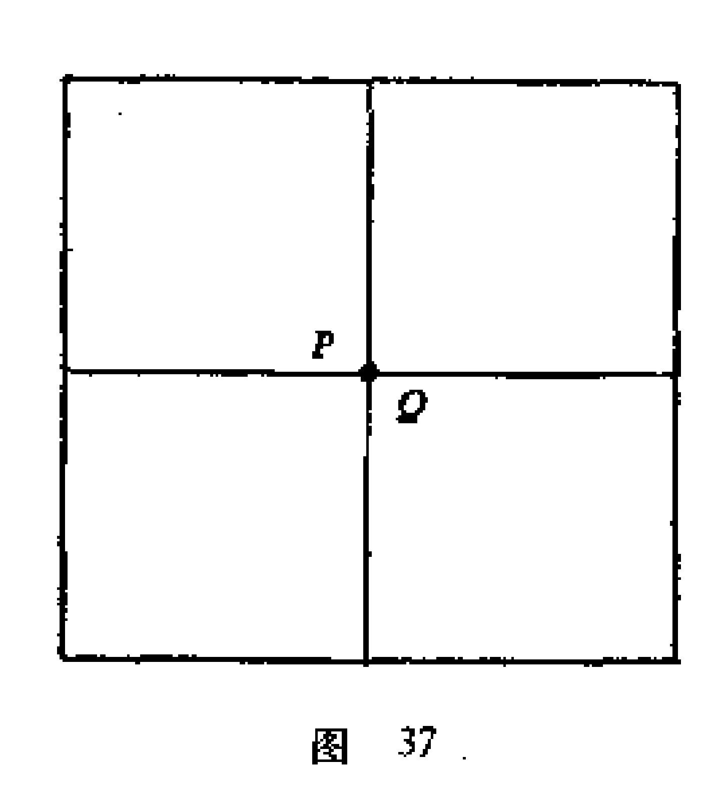


图 37

我们看到，应用这个方法做了第二批试验后，每做一批，区域的面积减少一半，因此收敛性是没有问题的，但有一个特殊情况要注意，就是会不会出现 P 点和 Q 点恰巧都是中点的情况，即横线纵线上的值都没有中心点的值大，这说明如果等高线

$$f(x, y)=d$$

不缩成一点(制高点)，则等高线在这点不存在切线，或者说切线在这点不连续。

(2) 有效性

如果单因素是用黄金分割法进行的，假定 x, y 所要求的精密度各为 ε 及 η，经适当的放大缩小，不妨假定 $\varepsilon=\eta$。在第一条线上作 l 次试验后，所达到的精密度是

$$(b-a)w^{l-1}, \quad w=\frac{1}{2}(-1+\sqrt{5}),$$

现在所要求的是

$$(b-a)w^{l-2}>\varepsilon \geqslant (b-a)w^{l-1},$$

即

$$\left(\frac{1}{w}\right)^{l-1} \geqslant \frac{b-a}{\varepsilon}>\left(\frac{1}{w}\right)^{l-2},$$

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

$$l-1 \geqslant \frac{\log \frac{b-a}{\varepsilon}}{\log \frac{1}{w}} > l-2,$$

也就是 l 差不多等于

$$\log \frac{b-a}{\varepsilon} \bigg/ \log \frac{1}{w} \sim \log \frac{1}{\varepsilon} \bigg/ \log \frac{1}{w}.$$

以后每做一批，区域面积去掉 1/2，经 p 次单变量处理后，得出的面积是原来的

$$\frac{1}{2^{p-1}}.$$

如果做到

$$\frac{(b-a)(d-c)}{2^{p-1}} \leqslant \varepsilon^2 < \frac{(b-a)(d-c)}{2^{p-2}},$$

当然就满足要求了，也就是 p 大致等于

$$\frac{2\log \frac{1}{\varepsilon}}{\log 2}.$$

因此，

$$2\left(\log \frac{1}{\varepsilon}\right)^2 \bigg/ \left(\log \frac{1}{\omega} \cdot \log 2\right)$$

次试验可达目的，也就是当 ε 很小时，无穷大的阶是

$$\left(\log \frac{1}{\varepsilon}\right)^2,$$

比起网格法 $\frac{1}{\varepsilon^2}$，要好得多.

附记 在 P，Q 二点做出来的试验数据如果相等(或无法辨认好坏)，这是件好事，因为这说明 P 点和 Q 点位于同一条等高线上了(见图 38)，所以可以把图上的下半块和左半块都去掉，仅留下第一象限.这一点对统计方法来说也是十分有利

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

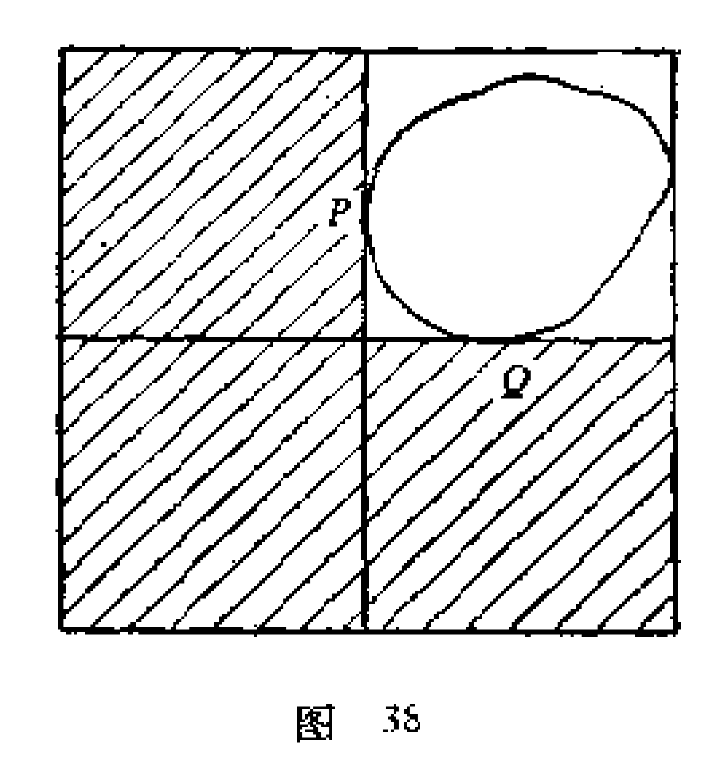


图 38

的，当两点试验数据的可分辨性十分接近时，我们干脆丢掉面积的 3/4。

§2. 旋 升 法

在进行旋升法的时候，请注意：划掉一个区域的原则，也就是有了 P_1 点与 P_2 点之后，划掉通过 P_1 点的直线所分开的不包含 P_2 点的部分，这样可以少做不少试验，当然，z 值不断上升的情况可能是外向的，也可能是内向的，也可能兼而有之（图 40）。

从方法来看，每次求最大值，可能比对开法好些，但并不

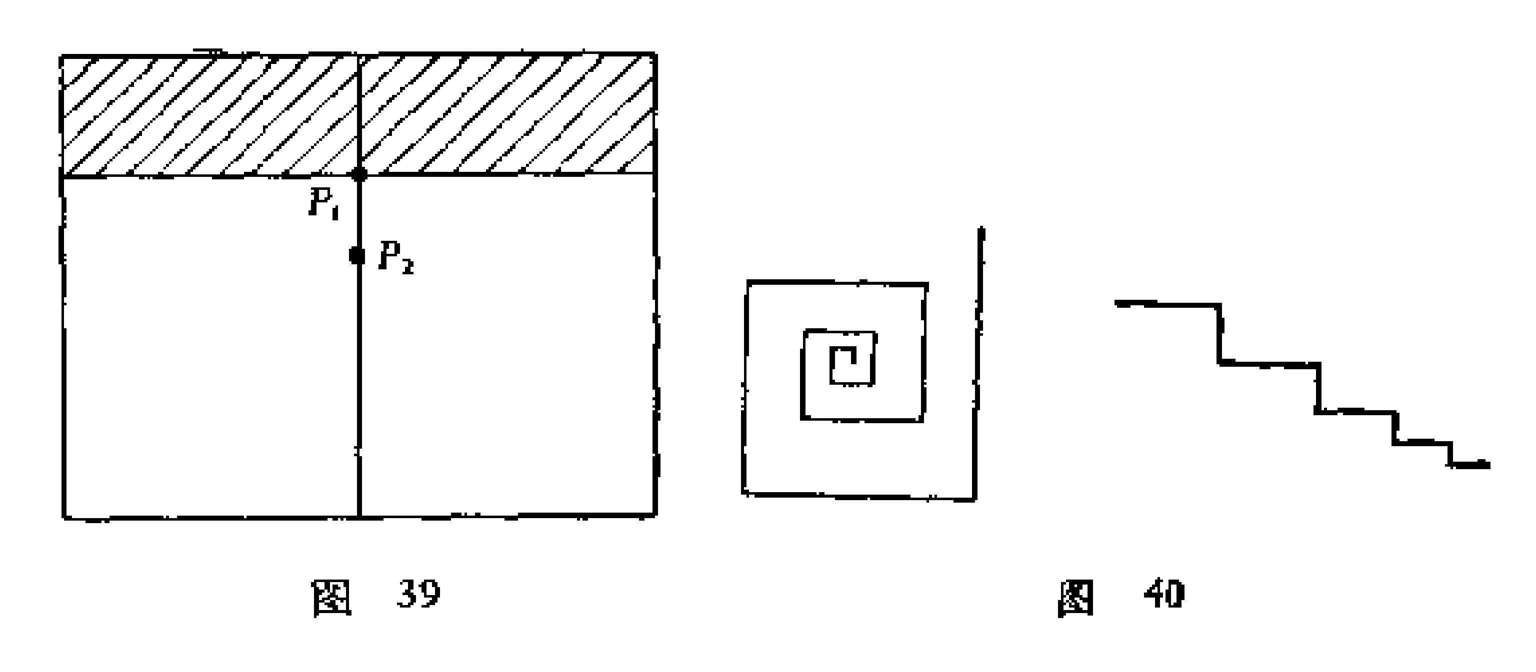


图 39　　　　图 40

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

能保证每次划去的部分大于或等于原来面积的一半．也就是说，上算的可能性虽然很大，但不上算的可能性也是不能排除的，不象对开法稳拿一半．

(1) 收敛性

如果一直做下去，则

$$z_1 = f(p_1) < z_2 = f(p_2) < \cdots < z_n = f(p_n) < \cdots$$

是一个单调上升的数列，而由问题性质看出这数列是有界的，因此

$$\lim_{n\to\infty} z_n = z_0 \quad (z_n \leqslant z_0).$$

我们画下高度等于

$$z_1, z_2, \cdots, z_n, \cdots$$

的等高线 $c_1, c_2, \cdots, c_n, \cdots$．如果 $z = z_0$ 不是制高点的高程，则有一条等高线 c_0，其高程是 z_0．当 n 充分大时，c_n 在 p_n 的法线将穿入 c_0 之内，换言之，z_{n+1} 比 z_0 更大，这是矛盾．p_l 在 c_{l-1} 之中，所以 p_l 趋于制高点．

(2) 有效性

要使误差小于 $\varepsilon(>0)$，需要的试验次数的数量级是

$$\left(\log \frac{1}{\varepsilon}\right)^2$$

证明请看下一章的§1.

§3. 平 行 线 法

从对开法的讨论中立刻明白平行线法一定逐步逼近最大值．如果单因素是用黄金分割法，则每条直线上要试验的次数大约是

$$\log \frac{1}{\varepsilon} \Big/ \log \frac{1}{w} \quad \left(w = \frac{\sqrt{5}-1}{2}\right),$$

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

而每做一条线上的试验后，区域面积缩小为原来的 w，这样 p 次后，留下的面积是原来的

$$w^{p-1}.$$

当要求 x 和 y 都达到 ε 的精度，p 的阶数约为

$$\log\frac{1}{\varepsilon}\Big/\log\frac{1}{w},$$

因此一共要做的试验次数约为

$$\left(\log\frac{1}{\varepsilon}\right)^2\Big/\left(\log\frac{1}{w}\right)^2.$$

同样可以讨论分批试验的情况.

§4. 两个因素的离散情形

例如优选的范围是一个 21 × 13 的格子，我们现在不用对开法，因为平分下来不一定在格子点上，也不一定上算，怎么办呢？我们在 $x=13$ 的直线上用分数法做五次试验．又在 $y=8$ 的直线上也用分数法，这时 T 点已做过试验，因此做 6—1 次试验．各得一最优点，记之为 P 点，Q 点．比较 P 点与 Q 点，如果 Q 点比 P 点好，则裁取一个 8 × 13 的格子(图 42)．

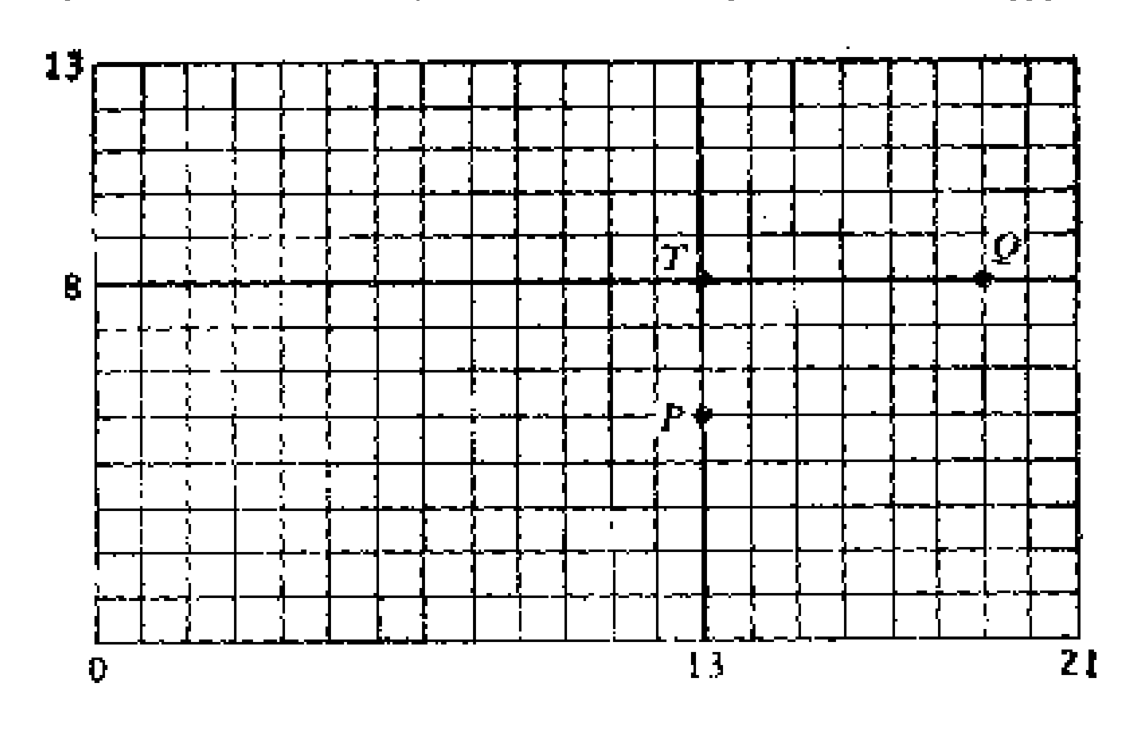


图 41

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

在这个图形中，取 $x=13+5=18$ 还是取 $x=21-5=16$，可以这样来考虑：因为 $x=18$ 更靠近 Q 点些，因而取 $x=18$，再在这条线上用分数法．这样做并不能排除不上算的可能，但上算的可能性大些(因为近"大"处得"更大")．

如果在每个格子点上作试验，共需试验 $20\times12=240$ 次，用现在的方法，最多只要 30 次就可以了，如果纵横格子个数并不恰巧等于某一 F_n，那末可以添上一些或冒点风险减少一些，以凑成 F_n．例如在

$$0<x<18$$

时，不妨添上些成 $0<x<21$，或减掉些，而成

$$0<x<13.$$

图 42

当然也可以用平行线法，读者自行补充．

§5. 翻筋斗法

当因素的变化不能跳跃式调整时，此方法有其优点，翻筋斗法的理论根据是：通过一点的两个方向的微分都等于零，一般来说，这点就是极值点．

其实前面关于等边三角形的限制不是必需的，根据具体情况用直角三角形，任意三角形都行．如果用四边形，五边形，…来翻筋斗效果如何呢？例如，有人建议用矩形，在矩形的四个顶点做试验，然后依最好点翻过去，在新的三个顶点做试验，等等，称为"矩形调优法"．这样做好不好？这样既多做了试验，又没有很好的理论根据．因为这里同时考虑了三个方向，而极值问题一般只要考虑两个方向就够了．因此"矩

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

形调优"就不如用三角形翻筋斗．翻筋斗法不难推广到多个因素的情形，例如：三个因素，用正四面体在四个顶点做试验，如图所示，依最好的一点翻过去，等等．

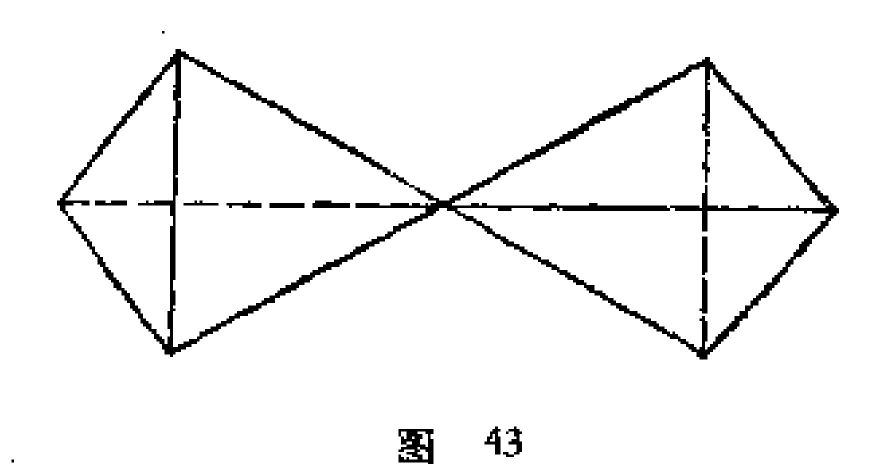


图 43

关于 n 个因素则是找出一个 n 维的正单纯形，在它的 $n+1$ 个顶点做试验，怎样找顶点？我们建议用以下的方法：

设 A 点的坐标是 $(a_1, a_2, \cdots, a_n)$，以 A 为顶点边长为 h 的 n 个顶点 $A_1, A_2, \cdots, A_n$ 的坐标可用下法得之，先做 B 点：

$$b_1 = a_1 + ah, \quad b_2 = a_2 + ah, \cdots, \quad b_n = a_n + ah,$$

这儿

$$a = \frac{1}{\sqrt{2}} \cdot \frac{\sqrt{n+1}-1}{n},$$

也就是每一坐标上都加上 ah，A_1 就是 B 的第一个坐标加上 $\frac{h}{\sqrt{2}}$，A_2 就是 B 的第二个坐标加上 $\frac{h}{\sqrt{2}}$，$\cdots$，A_n 就是 B 的第 n 个坐标加上 $\frac{h}{\sqrt{2}}$，其余坐标不变．证明是不难的，A 与 A_i 的距离平方等于：

$$(n-1)(ah)^2 + \left(ah + \frac{h}{\sqrt{2}}\right)^2$$

$$= \left[(n-1)a^2 + a^2 + \sqrt{2}a + \frac{1}{2}\right]h^2$$

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

$$= \left(na^2 + \sqrt{2}a + \frac{1}{2}\right)h^2$$

$$= \frac{1}{2}\left[\frac{(\sqrt{n+1}-1)^2}{n} + 2\frac{\sqrt{n+1}-1}{n} + 1\right]h^2$$

$$= h^2,$$

所以 A 与 A_i 的距离等于 h。

又 A_i 与 $A_j(j \neq i)$ 的距离等于

$$\sqrt{2\left(\frac{h}{\sqrt{2}}\right)^2} = h.$$

另一方法：假设 A 是原点，$h=1$（都不失其普遍性），取

$$A_1 = (1, 0, 0, \cdots, 0),$$

$$A_2 = (\beta_1, a_2, 0, \cdots, 0),$$

由于 A_2 与 A 的距离等于 1，A_1 与 A_2 的距离也等于 1，所以

$$\beta_1^2 + a_2^2 = 1, \quad (\beta_1 - 1)^2 + a_2^2 = 1.$$

得 $\beta_1 = \frac{1}{2}$，$a_2 = \pm\frac{\sqrt{3}}{2}$ 取正号得

$$A_2 = \left(\frac{1}{2}, \frac{\sqrt{3}}{2}, 0, \cdots, 0\right).$$

我们可以一维一维地添加，一般有

$$A_l = (\beta_1, \beta_2, \cdots, \beta_{l-1}, a_l, 0, \cdots, 0),$$

这儿

$$\beta_{l-1} = \sqrt{\frac{1}{2l(l+1)}}, \quad a_l = \sqrt{\frac{l+1}{2l}}.$$

证明：A 与 A_l 的距离平方等于

$$\beta_1^2 + \beta_2^2 + \cdots + \beta_{l-1}^2 + a_l^2$$

$$= \sum_{p=1}^{l-1} \frac{1}{2p(p+1)} + \frac{l+1}{2l}$$

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

$$= \sum_{p=1}^{l-1}\left(\frac{1}{2p} - \frac{1}{2(p+1)}\right) + \frac{l+1}{2l}$$

$$= \frac{1}{2} - \frac{1}{2l} + \frac{l+1}{2l} = 1.$$

又 A_l 与 $A_m(l > m)$ 的距离平方等于

$$(\alpha_m - \beta_m)^2 + \beta_{m+1}^2 + \cdots + \beta_{l-1}^2 + \alpha_l^2$$

$$= \left(\sqrt{\frac{m+1}{2m}} - \sqrt{\frac{1}{2m(m+1)}}\right)^2$$

$$+ \sum_{p=m+1}^{l-1} \frac{1}{2p(p+1)} + \frac{l+1}{2l}$$

$$= \frac{m}{2(m+1)} + \sum_{p=m+1}^{l-1}\left(\frac{1}{2p} - \frac{1}{2(p+1)}\right) + \frac{l+1}{2l}$$

$$= \frac{1}{2} - \frac{1}{2(m+1)} + \sum_{p=m+1}^{l-1}\left(\frac{1}{2p} - \frac{1}{2(p+1)}\right)$$

$$+ \frac{1}{2l} + \frac{1}{2} = 1.$$

即得所证，

新的试验点是很容易求出的： 在 $(n+1)$ 个顶点 P_1, $P_2, \cdots, P_{n+1}$ 上做试验，其中

$$P_i = (x_{i_1}, x_{i_2}, \cdots, x_{i_n}) \quad (i = 1, 2, \cdots, n+1).$$

假设 P_j 上的试验结果最好，依 P_j 找出各顶点的对称点

$$P_i'(i = 1, 2, \cdots, n+1):$$

$$P_i' - P_j = P_j - P_i,$$

即

$$P_i' = 2P_j - P_i, \quad (i = 1, 2, \cdots, n+1).$$

然后在新的顶点做试验，找出最好的一点，又依这点找对称点，一直做下去，如果做不下去了，比如 P_i 仍是最好点，这时将距离缩小，例如缩小一半：

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

$$P_i'' - P_i' = \frac{1}{2}(P_i' - P_j')$$

即

$$P_i'' = \frac{1}{2}(P_i' + P_j')$$

在 $P_1'', P_2'', \cdots, P_{n+1}''$ 继续做试验。

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

第二章 最陡上升法

述　　要

先在 $(n+1)$ 个点 $(x_1, x_2, \cdots x_n)$, $(x_1+\delta_1, x_2, \cdots, x_n)$, $\cdots(x_1, x_2, \cdots, x_{n-1}, x_n+\delta_n)$ 上各做一次试验，得出数据

$$z_1, z_1^{(1)}, \cdots, z_1^{(n)},$$

在线段

$$x_1+\frac{z_1^{(1)}-z_1}{\delta_1}t, \cdots, x_n+\frac{z_1^{(n)}-z_1}{\delta_n}t$$

上用单变数方法找出 z 的最大值，这儿 t 的变化范围由以下的不等式决定

(i) $$t \geqslant 0,$$

(ii) $$a_i \leqslant x_i+\frac{z_1^{(i)}-z_1}{\delta_i}t \leqslant b_i \quad (i=1, 2, \cdots, n).$$

如果找到的最大值在 $t=t_0$ 时取，则令

$$y_i = x_i+\frac{z_1^{(i)}-z_1}{\delta_i}t_0,$$

以 $(y_1, y_2, \cdots, y_n)$ 代替 $(x_1, x_2, \cdots, x_n)$ 继续进行，z 的值不断增大，一直做到 $(x_1, x_2, \cdots, x_n)$ 与 $(y_1, y_2, \cdots, y_n)$ 的距离小于预给的精密度为止。

§1. 最陡上升法

假定在区域 $a \leqslant x \leqslant b$, $c \leqslant y \leqslant d$ 内，

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

$$z = F(x, y) \tag{1}$$

是一个向上凸的曲面，处处有二阶连续微商，并且假定在讨论的区域内仅有唯一的制高点。

在制高点 (x_0, y_0) 处，

$$z = F(x, y) \tag{2}$$

一定适合

$$F_x(x_0, y_0) = 0, \quad F_y(x_0, y_0) = 0. \tag{3}$$

由于曲面向上凸，所以

$$\frac{\partial^2 F}{\partial x^2}\xi^2 + 2\frac{\partial^2 F}{\partial x \partial y}\xi\eta + \frac{\partial^2 F}{\partial y^2}\eta^2$$

是一个定负的二次型，也就是一定存在两个正数 M 与 m 使

$$-M(\xi^2+\eta^2) \leqslant \frac{\partial^2 F}{\partial x^2}\xi^2 + 2\frac{\partial^2 F}{\partial x \partial y}\xi\eta + \frac{\partial^2 F}{\partial y^2}\eta^2$$
$$\leqslant -m(\xi^2+\eta^2). \tag{4}$$

由闭区域内连续函数的性质，M，m 与 x，y 无关。

从一点 (x_1, y_1) 出发，若 $F(x_1, y_1) = z_1$ 对等高线

$$F(x, y) = z_1 \tag{5}$$

沿着法线方向作直线

$$x = x_1 + \left(\frac{\partial F}{\partial x}\right)_1 t, \quad y = y_1 + \left(\frac{\partial F}{\partial y}\right)_1 t \quad (t \geqslant 0). \tag{6}$$

用单变数法，沿此直线方向找

$$g(t) = F\left(x_1 + \left(\frac{\partial F}{\partial x}\right)_1 t, y_1 + \left(\frac{\partial F}{\partial y}\right)_1 t\right)$$

的最大值，假定在 $t = t_1$ 时取最大值，令

$$z_2 = g(t_1) = F\left(x_1 + \left(\frac{\partial F}{\partial x}\right)_1 t_1, y_1 + \left(\frac{\partial F}{\partial y}\right)_1 t_1\right) \tag{7}$$
$$= F(x_2, y_2),$$

由于 $g(t)$ 也是凸函数，因此

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

$$\frac{d}{dt}g(t)\Big|_{t=t_1}=\frac{d}{dt}F\left(x_1+\left(\frac{\partial F}{\partial x}\right)_1 t,\right.$$

$$\left.y_1+\left(\frac{\partial F}{\partial y}\right)_1 t\right)\Big|_{t=t_1}=0. \tag{8}$$

我们现在研究(与制高点高程差的进展比例)

$$\frac{z_0-z_2}{z_0-z_1}=\frac{F(x_0,y_0)-F(x_2,y_2)}{F(x_0,y_0)-F(x_1,y_1)}$$

$$=1-\frac{F(x_2,y_2)-F(x_1,y_1)}{F(x_0,y_0)-F(x_1,y_1)}. \tag{9}$$

1) 分母

$$F(x_0,y_0)-F(x_1,y_1)=-\int_0^1\frac{d}{dt}F(x_0+t(x_1-x_0),$$

$$y_0+t(y_1-y_0))dt=-\int_0^1dt\int_0^t\frac{d^2}{ds^2}F(x_0+s(x_1-x_0),$$

$$y_0+s(y_1-y_0)ds=-\int_0^1dt\int_0^t[(x_1-x_0)^2F_{x^2}$$

$$+2(x_1-x_0)(y_1-y_0)F_{xy}+(y_1-y_0)^2F_{y^2}]ds,$$

由(4)可知

$$F(x_0,y_0)-F(x_1,y_1)\leqslant M[(x_1-x_0)^2$$

$$+(y_1-y_0)^2]\int_0^1dt\int_0^t ds$$

$$=\frac{M}{2}[(x_1-x_0)^2+(y_1-y_0)^2]. \tag{10}$$

2) 再有(根据(7))

$$F(x_2,y_2)-F(x_1,y_1)=\int_0^{t_1}\frac{d}{dt}F\left(x_1+\left(\frac{\partial F}{\partial x}\right)_1 t,\right.$$

$$\left.y_1+\left(\frac{\partial F}{\partial y}\right)_1 t\right)dt=-\int_0^{t_1}dt\int_t^{t_1}\frac{d^2}{ds^2}F\left(x_1+\left(\frac{\partial F}{\partial x}\right)_1 s,\right.$$

$$\left.y_1+\left(\frac{\partial F}{\partial y}\right)_1 s\right)ds=-\int_0^{t_1}dt\int_t^{t_1}\left(\left(\frac{\partial F}{\partial x}\right)_1^2F_{x^2}\right.$$

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

$$+2\left(\frac{\partial F}{\partial x}\right)_1\left(\frac{\partial F}{\partial y}\right)_1 F_{xy}+\left(\frac{\partial F}{\partial y}\right)_1^2 F_{y^2}\Big)ds$$

$$\geqslant m\left[\left(\frac{\partial F}{\partial x}\right)_1^2+\left(\frac{\partial F}{\partial y}\right)_1^2\right]\int_0^{t_1}(t_1-t)dt$$

$$=\frac{m}{2}t_1^2\left[\left(\frac{\partial F}{\partial x}\right)_1^2+\left(\frac{\partial F}{\partial y}\right)_1^2\right]. \tag{11}$$

3) 从 (8) 式推得

$$F_x\left[x_1+\left(\frac{\partial F}{\partial x}\right)_1 t_1,\ y_1+\left(\frac{\partial F}{\partial y}\right)_1 t_1\right]\left(\frac{\partial F}{\partial x}\right)_1+F_y\left[x_1+\left(\frac{\partial F}{\partial x}\right)_1 t_1,\ y_1+\left(\frac{\partial F}{\partial y}\right)_1 t_1\right]\left(\frac{\partial F}{\partial y}\right)_1=0 \tag{12}$$

我们把

$$F_x\left(x_1+\left(\frac{\partial F}{\partial x}\right)_1 t_1,\ y_1+\left(\frac{\partial F}{\partial y}\right)_1 t_1\right)-F_x(x_1,\ y_1)$$

$$=\int_0^{t_1}\frac{d}{dt}F_x\left(x_1+\left(\frac{\partial F}{\partial x}\right)_1 t,\ y_1+\left(\frac{\partial F}{\partial y}\right)_1 t\right)dt$$

及

$$F_y\left(x_1+\left(\frac{\partial F}{\partial x}\right)_1 t_1,\ y_1+\left(\frac{\partial F}{\partial y}\right)_1 t_1\right)-F_y(x_1,\ y_1)$$

$$=\int_0^{t_1}\frac{d}{dt}F_y\left(x_1+\left(\frac{\partial F}{\partial x}\right)_1 t,\ y_1+\left(\frac{\partial F}{\partial y}\right)_1 t\right)dt$$

代入 (12) 可知

$$\left(\frac{\partial F}{\partial x}\right)_1^2+\left(\frac{\partial F}{\partial y}\right)_1^2=-\int_0^{t_1}\left[\left(\frac{\partial F}{\partial x}\right)_1\frac{d}{dt}F_x\left(x_1+\left(\frac{\partial F}{\partial x}\right)_1 t, y_1+\left(\frac{\partial F}{\partial y}\right)_1 t\right)+\left(\frac{\partial F}{\partial y}\right)_1\frac{d}{dt}F_y\left(x_1+\left(\frac{\partial F}{\partial x}\right)_1 t,\ y_1+\left(\frac{\partial F}{\partial y}\right)_1 t\right)\right]dt=-\int_0^{t_1}\left[\left(\frac{\partial F}{\partial x}\right)_1^2 F_{x^2}+2\left(\frac{\partial F}{\partial x}\right)_1\left(\frac{\partial F}{\partial y}\right)_1 F_{xy}+\left(\frac{\partial F}{\partial y}\right)_1^2 F_{y^2}\right]dt, \tag{13}$$

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

由（4）可知

$$m\left[\left(\frac{\partial F}{\partial x}\right)_1^2+\left(\frac{\partial F}{\partial y}\right)_1^2\right]t_1\leqslant\left(\frac{\partial F}{\partial x}\right)_1^2+\left(\frac{\partial F}{\partial y}\right)_1^2$$
$$\leqslant M\left[\left(\frac{\partial F}{\partial x}\right)_1^2+\left(\frac{\partial F}{\partial y}\right)_1^2\right]t_1,$$

即得

$$\frac{1}{M}\leqslant t_1\leqslant\frac{1}{m}. \tag{14}$$

4）由（3）可知

$$\left(\frac{\partial F}{\partial x}\right)_1=F_x(x_1, y_1)-F_x(x_0, y_0)$$
$$=\int_0^1\frac{d}{dt}F_x(x_0+t(x_1-x_0), y_0+t(y_1-y_0))dt$$

及

$$\left(\frac{\partial F}{\partial y}\right)_1=\int_0^1\frac{d}{dt}F_y(x_0+t(x_1-x_0), y_0+t(y-y_0))dt,$$

并推出

$$\left(\frac{\partial F}{\partial x}\right)_1(x_1-x_0)+\left(\frac{\partial F}{\partial y}\right)_1(y_1-y_0)=\int_0^1[(x_1-x_0)^2F_{x^2}+2(x_1-x_0)(y_1-y_0)F_{xy}+(y_1-y_0)^2F_{y^2}]dt$$

由于 $\xi^2F_{x^2}+2\xi\eta F_{xy}+\eta^2F_{y^2}$ 是定负的，所以

$$\left|\left(\frac{\partial F}{\partial x}\right)_1(x-x_0)+\left(\frac{\partial F}{\partial y}\right)_1(y_1-y_0)\right|=\int_0^1|(x_1-x_0)^2F_{x^2}+2(x_1-x_0)(y_1-y_0)F_{xy}+(y_1-y_0)^2F_{y^2}|dt\geqslant m[(x_1-x_0)^2+(y_1-y_0)^2]. \tag{15}$$

利用 Schwarz 不等式得到

$$m[(x_1-x_0)^2+(y_1-y_0)^2]\leqslant\left[\left(\frac{\partial F}{\partial x}\right)_1^2\right.$$

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

$$+\left(\frac{\partial F}{\partial y}\right)_1^2\Bigg]^{\frac{1}{2}}[(x-x_0)^2+(y_1-y_0)^2]^{\frac{1}{2}},$$

于是

$$m^2[(x_1-x_0)^2+(y_1-y_0)^2]\leqslant\left(\frac{\partial F}{\partial x}\right)_1^2+\left(\frac{\partial F}{\partial y}\right)_1^2. \quad (16)$$

综合 1)，2)，3) 及 4) 的结果，可知

$$\frac{F(x_2,y_2)-F(x_1,y_1)}{F(x_0,y_0)-F(x_1,y_1)}\geqslant\frac{\frac{1}{2}mt_1^2\left[\left(\frac{\partial F}{\partial x}\right)_1^2+\left(\frac{\partial F}{\partial y}\right)_1^2\right]}{\frac{1}{2}M[(x_1-x_0)^2+(y_1-y_0)^2]}$$

$$\geqslant\frac{m}{M^3}\frac{\left(\frac{\partial F}{\partial x}\right)_1^2+\left(\frac{\partial F}{\partial y}\right)_1^2}{(x_1-x_0)^2+(y_1-y_0)^2}$$

$$\geqslant\frac{m^3}{M^3},$$

也就是

$$\frac{z_0-z_2}{z_0-z_1}\leqslant 1-\left(\frac{m}{M}\right)^3. \quad (17)$$

这个结果值得注意的地方是：右边与 z_1 无关，也就是从任一点开始，做一次后，与制高点的高程差缩小 $1-\left(\frac{m}{M}\right)^3$ 倍，如果做了 l 次，则高程差与原来的高程差的比例是

$$\left(1-\left(\frac{m}{M}\right)^3\right)^l.$$

取 l 使

$$\left[1-\left(\frac{m}{M}\right)^3\right]^l\leqslant\varepsilon<\left[1-\left(\frac{m}{M}\right)^3\right]^{l-1},$$

则必

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

$$l \geqslant \frac{\log \frac{1}{\varepsilon}}{\log \frac{1}{1-\left(\frac{m}{M}\right)^{3}}} > l-1,$$

也就是 l 的数量级等于 $\log \frac{1}{\varepsilon}$.

在一条直线上做黄金分割法的数量级也是 $\log \frac{1}{\varepsilon}$，而每找一陡升方向需要做二个试验，因此总共的试验次数的数量级是

$$\left(\log \frac{1}{\varepsilon}\right)^{2}.$$

这方法可以推广到 s 个因素，而由于每找一次陡升方向需做 s 个试验，所以总的试验次数的数量级仍是

$$\left(\log \frac{1}{\varepsilon}\right)^{2}.$$

最陡上升法形式上与第一章所讲的旋升法有所不同，实质上是一样的．旋升法的要点是先求一条直线 l_1 上 $F(x, y)$ 的最大值，取这最大值的点 (x_1, y_1) 就是这直线与某等高线的切点，陡升方向就是法线方向，也就是在 (x_1, y_1) 点找过这点的等高线的法线方向，因而原则上并无不同.

旋升法不但易行，而且不必要用差分逼近微分，不必要每次做二个试验来确定陡升方向.

§2. 渐近陡升法

在上节所讨论的最陡上升法中，如果函数 $F(x, y)$ 的表达式已知，二个偏导数 $\frac{\partial F}{\partial x}$, $\frac{\partial F}{\partial y}$ 可以通过计算直接得到．但在实际应用时，对所求的目标函数绝大部分是连续函数的形

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

式，是不知道的．因此，要求出在一点的陡升方向也只有用做试验的办法，一般可以这样做：在点(x, y)附近，沿坐标轴的方向各找一点$(x+\delta_1, y)$和$(x, y+\delta_2)$，在这二点做试验，并计算出

$$\frac{f(x+\delta_1, y)-f(x, y)}{\delta_1} \text{和} \frac{f(x, y+\delta_2)-f(x, y)}{\delta_2}.$$

当δ_1，δ_2取得适当小时，它们便给出点(x, y)处的两个偏导数的（近似）值．

把这个方法过渡到一般的情况，便是渐近陡升法了．

保持本章开始时述要部分所用的记号．

我们现在每次沿最陡的方向上升距离为ε的一步，也就是取t使

$$x_1+\frac{z_1^{(1)}-z_1}{\delta_1}t, \cdots, x_n+\frac{z_1^{(n)}-z_1}{\delta_n}t \quad (t>0)$$

与$(x_1, \cdots, x_n)$的距离等于

$$t\sqrt{\left(\frac{z_1^{(1)}-z_1}{\delta_1}\right)^2+\cdots+\left(\frac{z_1^{(n)}-z_1}{\delta_n}\right)^2}=\varepsilon.$$

取ε小到可以分辨的情况，则得渐近陡升法，这就是在山坡上，每次向最陡的方向走上长度为ε的一步，这个办法虽然不如前法快——一步跨到山脊——但是对于不能跳跃做试验的情况比较合适．

一般讲来，我们可以取合适的单位使

$$\delta_1=\delta_2=\cdots=\delta_n,$$

并且不妨假定$\delta_i=\varepsilon$，也就是用

$$x_i+\frac{z_1^{(i)}-z_1}{\sqrt{(z_1^{(1)}-z_1)^2+\cdots+(z_1^{(n)}-z_1)^2}}\varepsilon$$
$$(i=1, 2, \cdots, n)$$

来代替原来的$(x_1, x_2, \cdots, x_n)$．

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

§3. 二次模拟

在 §1 中我们找到收敛因子

$$\left(1-\left(\frac{m}{M}\right)^{3}\right).$$

在统计学中,我们经常(不太令人放心地)假定数学模型是二次的.下面将证明,如果真是二次的,收敛因子还可以有所改进,这说明在靠近最优点时,也就是用 Taylor 级数的二次及二次以下的项来逼近时,实际上收敛速度要比 §1 所讲的要快得多.

也就是我们检查最陡上升法用在以下特例上的收敛情况:

$$F(x_1, \cdots, x_n) = c + 2\sum_{i=1}^{n} b_i x_i - \sum_{i,j=1}^{n} a_{ij}x_i x_j,$$

这儿 $a_{ij}=a_{ji}$, (a_{ij}) 是定正的对称方阵,它的特征根

$$0<\lambda_1 \leqslant \lambda_2 \leqslant \cdots \leqslant \lambda_n.$$

经过刚体运动,以及在 F 上减一常数,我们可以不失普遍性地考虑以下的特例:

$$Q(x_1, \cdots, x_n) = -\sum_{i=1}^{n} \lambda_i x_i^2,$$

它的最大值是 0, 在 $x_1=\cdots=x_n=0$ 处取.

从一定点 $(x_1^{(0)}, \cdots, x_n^{(0)})$ 出发,在这点的陡升方向是

$$(-2\lambda_1 x_1^{(0)}, \cdots, -2\lambda_n x_n^{(0)}).$$

研究

$$f(t) = Q(x_1^{(0)} - 2\lambda_1 x_1^{(0)}t, \cdots, x_n^{(0)} - 2\lambda_n x_n^{(0)}t)$$

的最大值,由于

$$f(t) = -\sum_{\nu=1}^{n} \lambda_\nu x_\nu^{(0)2}(1-2\lambda_\nu t)^2 = -\sum_{\nu=1}^{n} \lambda_\nu x_\nu^{(0)2}$$

$$+4t\sum_{\nu=1}^{n} \lambda_\nu^2 x_\nu^{(0)2} - 4t^2\sum_{\nu=1}^{n} \lambda_\nu^3 x_\nu^{(0)2} = -\sum_{\nu=1}^{n} \lambda_\nu x_\nu^{(0)2}$$

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

$$+\frac{\left(\sum_{\nu=1}^{n}\lambda_\nu^2 x_\nu^{(0)^2}\right)^2}{\sum_{\nu=1}^{n}\lambda_\nu^3 x_\nu^{(0)^2}}-\sum\lambda_\nu^3 x_\nu^{(0)^2}\left(2t-\frac{\sum\lambda_\nu^2 x_\nu^{(0)^2}}{\sum\lambda_\nu^3 x_\nu^{(0)^2}}\right)^2$$

所以，当

$$t=t^*=\frac{\sum\lambda_\nu^2 x_\nu^{(0)^2}}{2\sum\lambda_\nu^3 x_\nu^{(0)^2}}$$

时，即当

$$x_\nu^{(1)}=x_\nu^{(0)}-\lambda_\nu x_\nu^{(0)}\cdot\frac{\sum\lambda_\nu^2 x_\nu^{(0)^2}}{\sum\lambda_\nu^3 x_\nu^{(0)^2}}$$

时，取最大值

$$Q(x_1^{(1)},\cdots,x_n^{(1)})=f(t)^*=-\sum_{\nu=1}^{n}\lambda_\nu x_\nu^{(0)^2}$$

$$+\frac{\left(\sum_{\nu=1}^{n}\lambda_\nu^2 x_\nu^{(0)^2}\right)^2}{\sum_{\nu=1}^{n}\lambda_\nu^3 x_\nu^{(0)^2}}.$$

现在利用不等式(其证明在下节)

$$\sum_{\nu=1}^{n}\frac{a_\nu}{\lambda_\nu}\sum_{\nu=1}^{n}a_\nu\lambda_\nu\leqslant\frac{1}{4}\left(\sqrt{\frac{\lambda_1}{\lambda_n}}+\sqrt{\frac{\lambda_n}{\lambda_1}}\right)^2\cdot\left(\sum_{\nu=1}^{n}a_\nu\right)^2,$$

得

$$\frac{Q(x_1^{(1)},\cdots,x_n^{(1)})}{Q(x_1^{(0)},\cdots,x_n^{(0)})}=1-\frac{\left(\sum_{\nu=1}^{n}\lambda_\nu^2 x_\nu^{(0)^2}\right)^2}{\sum_{\nu=1}^{n}\lambda_\nu x_\nu^{(0)^2}\sum_{\nu=1}^{n}\lambda_\nu^3 x_\nu^{(0)^2}}$$

$$\leqslant 1-\frac{4}{\left(\sqrt{\frac{\lambda_1}{\lambda_n}}+\sqrt{\frac{\lambda_n}{\lambda_1}}\right)^2}$$

$$=\left(\frac{\lambda_n-\lambda_1}{\lambda_n+\lambda_1}\right)^2.$$

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

回到一般的形式，λ_n 就是 M，λ_1 就是 m，因此收敛因子是

$$\left(\frac{M-m}{M+m}\right)^2,$$

且

$$\left(\frac{M-m}{M+m}\right)^2<\left(1-\left(\frac{m}{M}\right)^3\right).$$

由于下节末段所指出的论点，这因子是不能再改进的．这结果指出 §1 的收敛因子是有改进的可能性的，如果能找到一个结果连偏微商的差分逼近的误差也包括在内，那就更好．

§4. 反向 Schwarz 不等式

现在证明上节的不等式：如果

$$0<\lambda_1\leqslant\lambda_2\leqslant\cdots\leqslant\lambda_n,\quad a_i\geqslant 0,$$

则

$$\sum_{i=1}^{n}\lambda_i a_i\sum_{i=1}^{n}\frac{a_i}{\lambda_i}\leqslant\frac{1}{4}\left(\sqrt{\frac{\lambda_1}{\lambda_n}}+\sqrt{\frac{\lambda_n}{\lambda_1}}\right)^2\left(\sum_{i=1}^{n}a_i\right)^2.$$

证：此不等式的左边是 λ_i 的函数

$$f(\lambda_i)=p+q\lambda_i+\frac{r}{\lambda_j}\quad(p>0,\ q>0,\ r>0),$$

其二阶微商大于零，因此在 $\lambda_1<\lambda_i<\lambda_n$ 中不可能有最大值．仅当 $\lambda_i=\lambda_1$ 或 λ_n 时，它才能最大，也就是

$$\sum_{i=1}^{n}\lambda_i a_i\sum_{i=1}^{n}\frac{a_i}{\lambda_i}\leqslant\left(\lambda_1\sum_{i=1}^{m}a_i+\lambda_n\sum_{j=m+1}^{n}a_j\right)\left(\frac{1}{\lambda_1}\sum_{i=1}^{m}a_i\right.$$

$$\left.+\frac{1}{\lambda_n}\sum_{j=m+1}^{n}a_j\right)=\left(\sum_{i=1}^{m}a_i\right)^2+\left(\sum_{j=m+1}^{n}a_j\right)^2$$

$$+\left(\frac{\lambda_1}{\lambda_n}+\frac{\lambda_n}{\lambda_1}\right)\sum_{i=1}^{m}a_i\sum_{j=m+1}^{n}a_j=\left(\sum_{i=1}^{n}a_i\right)^2$$

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

$$+\left(\frac{\lambda_1}{\lambda_n}+\frac{\lambda_n}{\lambda_1}-2\right)\sum_{i=1}^{m}a_i\sum_{j=m+1}^{n}a_j\leqslant\left(\sum_{i=1}^{n}a_i\right)^2$$

$$+\left(\sqrt{\frac{\lambda_1}{\lambda_n}}-\sqrt{\frac{\lambda_n}{\lambda_1}}\right)^2\frac{\left(\sum\limits_{i=1}^{n}a_i\right)^2}{4}=\frac{1}{4}\left(\sqrt{\frac{\lambda_1}{\lambda_n}}\right.$$

$$\left.+\sqrt{\frac{\lambda_n}{\lambda_1}}\right)^2\left(\sum_{i=1}^{n}a_i\right)^2,$$

当且仅当

$$\lambda_1=\cdots=\lambda_m,\quad \lambda_{m+1}=\cdots=\lambda_n$$

及

$$\sum_{i=1}^{m}a_i=\sum_{j=m+1}^{n}a_j$$

时取等号。

§5. 收敛因子的进一步改进

在§1中我们粗略地估计出了

$$\frac{z_0-z_2}{z_0-z_1}\leqslant 1-\left(\frac{m}{M}\right)^3. \tag{1}$$

而在 §2 中，我们对二次目标函数证明了：收敛因子 $1-(m/M)^3$ 可以改进为 $\left(\frac{M-m}{M+m}\right)^2$. 在本节中我们将对一般的目标函数证明与二次目标函数相同的结果。

考虑

$$z=f(x), \tag{2}$$

这里 $x=(x_1, x_2, \cdots, x_n)$. 假设 $f(x)$ 具有三阶连续偏微商，且是一个一致的凹函数(在一维和二维时，即以前讲的“向上凸”函数)。

引进记号:

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

$$g(x) = \nabla f(x) = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \cdots, \frac{\partial f}{\partial x_s}\right) \tag{3}$$

称为 f 在点 x 处的梯度.

$$H(x) = \nabla^2 f(x) = \left(\frac{\partial^2 f(x)}{\partial x_i \partial x_j}\right)_{1\leqslant i,j\leqslant s} \tag{4}$$

称为 f 在点 x 的二阶微分矩阵,又称 Hessian 阵.

最陡上升法是从一点 $x^{(\nu)} = (x_1^{(\nu)}, \cdots, x_s^{(\nu)})$ 出发沿着梯度方向作直线

$$x = x^{(\nu)} + tg_\nu \quad (t \geqslant 0) \tag{5}$$

$(g_\nu = g(x^{(\nu)}))$, 用单变数法求出 $f(x)$ 在这直线上的最大值. 假定在 $t = t_\nu$ 时取此值,令

$$x^{(\nu+1)} = x^{(\nu)} + t_\nu g_\nu. \tag{6}$$

然后从 $x^{(\nu+1)}$ 出发重复上面的步骤迭代进行.

用 §1 同样的方法,可以估出

$$\begin{aligned} f(x^{(\nu)}) - f(x^{(0)}) &= \int_0^1 g(x^{(0)} + t(x^{(\nu)} - x^{(0)}))(x^{(\nu)} \\ &\quad - x^{(0)})' dt = \int_0^1 [g(x^{(0)} + t(x^{(\nu)} - x^{(0)})) \\ &\quad - g(x^{(0)})](x^{(\nu)} - x^{(0)})' dt = \int_0^1 dt \int_0^t (x^{(\nu)} \\ &\quad - x^{(0)})H(x^{(0)} + s(x^{(\nu)} - x^{(0)}))(x^{(\nu)} - x^{(0)})' ds \\ &= \int_0^1 dt \int_0^t (x^{(\nu)} - x^{(0)})(H_\nu + o(|x^{(\nu)} - x^{(0)}|))(x^{(\nu)} \\ &\quad - x^{(0)})' ds = \frac{1}{2}(x^{(\nu)} - x^{(0)})H_\nu(x^{(\nu)} - x^{(0)})' \\ &\quad + o(|x^{(\nu)} - x^{(0)}|^3). \end{aligned} \tag{7}$$

这里 $H_\nu = H(x^{(\nu)})$, $x^{(0)}$ 是 $z = f(x)$ 取最大值的点, 故有 $g(x^{(0)}) = 0$. 同样可得

$$\begin{aligned} g_\nu &= \int_0^1 (x^{(\nu)} - x^{(0)})H(x^{(0)} + t(x^{(\nu)} - x^{(0)}))dt \\ &= (x^{(\nu)} - x^{(0)})H_\nu + o(|x^{(\nu)} - x^{(0)}|^2). \end{aligned} \tag{8}$$

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

所以

$$x^{(\nu)} - x^{(0)} = g_\nu H_\nu^{-1} + o(|x^{(\nu)} - x^{(0)}|^2),$$

因此

$$f(x^{(\nu)}) - f(x^{(0)}) = \frac{1}{2} g_\nu H_\nu^{-1} g_\nu' + o(|x^{(\nu)} - x^{(0)}|^3). \tag{9}$$

又有

$$\begin{aligned} f(x^{(\nu)}) - f(x^{(\nu+1)}) &= -\int_0^{t_\nu} g(x^{(\nu)} + t g_\nu) g_\nu' dt \\ &= \int_0^{t_\nu} dt \int_t^{t_\nu} g_\nu H(x^{(\nu)} + s g_\nu) g_\nu' ds \\ &= \int_0^{t_\nu} dt \int_t^{t_\nu} g_\nu [H_\nu + o(|g_\nu|)] g_\nu' ds \\ &= \frac{1}{2} t_\nu^2 g_\nu H_\nu g_\nu' + o(|g_\nu|^3). \end{aligned} \tag{10}$$

由于 $f(x^{(\nu)} + t g_\nu)$ 在 $t = t_\nu$ 取极值，所以

$$\begin{aligned} 0 &= \frac{d}{dt} f(x^{(\nu)} + t g_\nu)\Big|_{t=t_\nu} = g(x^{(\nu)} + t_\nu g_\nu) g_\nu' \\ &= [g_\nu + t_\nu g_\nu H_\nu + o(|g_\nu|^2)] g_\nu', \end{aligned} \tag{11}$$

即

$$t_\nu = -\frac{g_\nu g_\nu'}{g_\nu H_\nu g_\nu'} + o(|g_\nu|). \tag{12}$$

代入(10)，得

$$f(x^{(\nu)}) - f(x^{(\nu+1)}) = \frac{1}{2} \frac{|g_\nu|^4}{g_\nu H_\nu g_\nu'} + o(|g_\nu|^3). \tag{13}$$

综合(9)和(13)，

$$\begin{aligned} \frac{f(x^{(0)}) - f(x^{(\nu+1)})}{f(x^{(0)}) - f(x^{(\nu)})} &= 1 - \frac{f(x^{(\nu+1)}) - f(x^{(\nu)})}{f(x^{(0)}) - f(x^{(\nu)})} \\ &= 1 - \frac{|g_\nu|^4 / g_\nu H_\nu g_\nu' + o(|g_\nu|^3)}{g_\nu H_\nu^{-1} g_\nu + o(|x^{(\nu)} - x^{(0)}|)^3} \\ &= 1 - \frac{|g_\nu|^4}{(g_\nu H_\nu^{-1} g_\nu')(g_\nu H_\nu g_\nu')} + o(|g_\nu|) \end{aligned} \tag{14}$$

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

根据假定，$f(x)$ 是一致凹的，所以 H_ν 定负，记其特征根为 $-\lambda_1, -\lambda_2, \cdots, -\lambda_s$, $(0 < \lambda_1 \leqslant \lambda_2 \leqslant \cdots \leqslant \lambda_s)$. 则有正交变换矩阵 Q，使

$$QH_\nu Q' = \begin{pmatrix} -\lambda_1 & & & 0 \\ & -\lambda_2 & & \\ & & \ddots & \\ 0 & & & -\lambda_s \end{pmatrix}.$$

记 $a_\nu = g_\nu Q$，(14) 式变为

$$\frac{f(x^{(0)}) - f(x^{(\nu+1)})}{f(x^{(0)}) - f(x^{(\nu)})} = 1 - \frac{|a_\nu|^4}{\sum_{k=1}^{s} \lambda_k a_{\nu,k}^2 \sum_{k=1}^{s} \frac{1}{\lambda_k} a_{\nu,k}^2}$$

$$+ o(|g_\nu|).$$

略去无穷小项，并应用反向 Schwarz 不等式得

$$\frac{f(x^{(0)}) - f(x^{(\nu+1)})}{f(x^{(0)}) - f(x^{(\nu)})} \leqslant 1 - \frac{4}{\left(\sqrt{\frac{\lambda_1}{\lambda_s}} + \sqrt{\frac{\lambda_s}{\lambda_1}}\right)^2}$$

$$= \left(\frac{\lambda_s - \lambda_1}{\lambda_s + \lambda_1}\right)^2.$$

于是我们得出

定理 若 $f(x) \in c^3$，且有

$$-Mgg' \leqslant gH(x)g' \leqslant -mgg',$$

则

$$\frac{z^{(0)} - z^{(\nu+1)}}{z^{(0)} - z^{(\nu)}} = \frac{f(x^{(0)}) - f(x^{(\nu+1)})}{f(x^{(0)}) - f(x^{(\nu)})} \leqslant \left(\frac{M - m}{M + m}\right)^2.$$

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

第三章　切　块　法

述　　要

我们的目的是要在区域 R_0 中找出函数

$$z = f(x, y)$$

的最大值，在 R_0 中取一点 $p_0(x_0, y_0)$，在 (x_0, y_0) 及其邻近两点$(x_0 + h, y_0)$，$(x_0, y_0 + k)$ 作试验，得 $f(x_0, y_0)$，$f(x_0 + h, y_0)$，$f(x_0, y_0 + k)$ 的数值，这里 h, k 以小为好，但又要使诸 f 值可以分辨．作

$$p_0': \left(x_0 + \frac{f(x_0 + h, y_0) - f(x_0, y_0)}{h},\right.$$

$$\left. y_0 + \frac{f(x_0, y_0 + k) - f(x_0, y_0)}{k}\right).$$

过 p_0 作垂直于 p_0，p_0' 联线的垂线 l_0，这垂线分 R_0 为两部分，丢掉不含有 p_0' 的部分，留下的部分用 R_1 表示，再在其中取一点 p_1，$\cdots$．

这方法的道理是显然的，因为分割是沿通过 p_0 的等高线的切线，且留下部分包括升高方向的一边．问题在于怎样选 p_0? 我们有下面几个建议．

第一个建议是：取 p_0 为 R_0 的重心，以下将证明，这一方法是有保证的．下文将证明，通过重心的任一直线把凸域分为两份，其面积的比在 4/5 及 5/4 之间，因此，每次留下的部分的面积不大于原面积的 5/9，做 n 次后，留下的部分不大于原面积乘 $(5/9)^n$，因而 n 充分大时，这方法保证是可以无限

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

精密的.

第二个建议是：取余下的部分中的法线的中点 p_1，作为下一次试验的 p_0，这方法简单易行，但还没有数学保证，以下的附记可供参考. 如果 p_1 做出来的实验比 p_0 小得多，可以不必在 p_1 附近再做试验，而在 p_1，p_0 的联线的中点 p_2 做试验. 如果还小得多，可以在 p_2，p_0 的中点 p_3 做试验，等等，得到和 p_0 所做出的差不多好或更好些的点 p^*，再在 p^* 用上面的方法，这样可以加强这个方法的可靠性.

第三个方法. 假设我们所研究的对象是多边形，其顶点各为 (x_i, y_i) $(i = 1, 2, \cdots, l)$，我们可以用

$$\left(\frac{1}{l}(x_1 + \cdots + x_l), \frac{1}{l}(y_1 + \cdots + y_l)\right)$$

作为下一个试点. 但通过这个点的直线，并没有分图形为两部其比例有上下限的性质.

当推广到 s 个变数时，找重心也并不容易，但本章的方法并不是没有实用价值的，而理论方面还有深入探讨的必要.

§1. 一个几何不等式(二维)

定理 通过一凸区域的重心的任意直线把凸域一分为二，其面积之比不小于 4/5，不大于 5/4.

为了替高维作准备，我们在讲几何直观证明的同时，还讲易于推广的解析形式.

凸域 D 的定义是：如果两点 (x_1, y_1)，(x_2, y_2) 在 D 内，则这两点间的直线段也在其中，即它含有

$$x = x_1 + t(x_2 - x_1), \quad y = y_1 + t(y_2 - y_1) \qquad (0 \leqslant t \leqslant 1). \tag{1}$$

区域 D 的面积等于

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

$$A = \iint_D dxdy, \tag{2}$$

其重心的坐标是

$$\bar{x} = \frac{1}{A}\iint_D xdxdy, \tag{3}$$

$$\bar{y} = \frac{1}{A}\iint_D ydxdy. \tag{4}$$

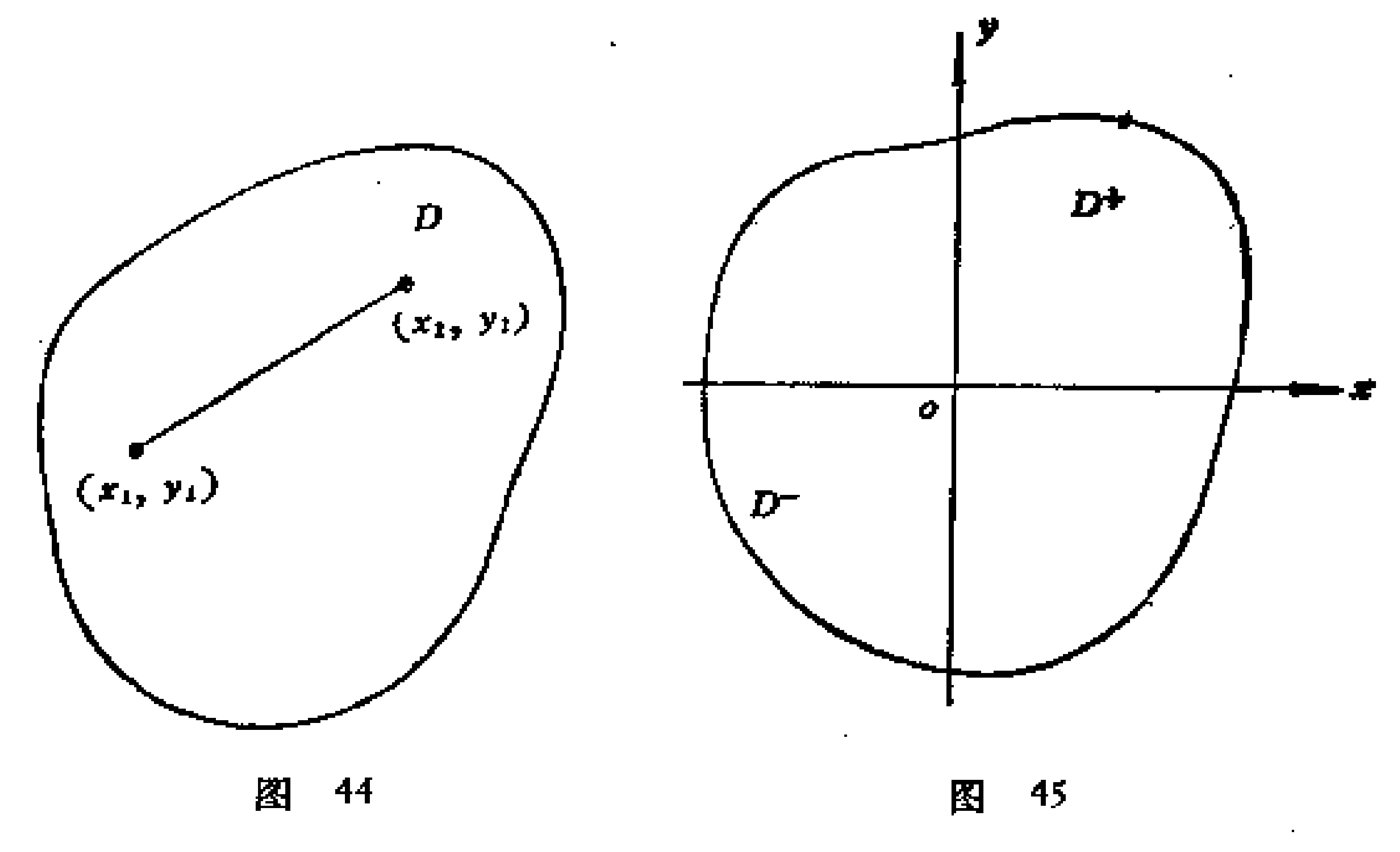


图 44　　　　图 45

假定分割D的直线就是x轴，则重心在x轴上，即

$$\iint_D ydxdy = 0, \tag{5}$$

x轴分D为两部分，轴上的用D^+表示，轴下的用D^-表示，并用$|D|$表示D的面积，则定理的结论就是

$$\left|\frac{D^+}{D^-}\right| \geqslant \frac{4}{5}, \quad \left|\iint_{D^+} dxdy\right| \geqslant \frac{4}{5}\left|\iint_{D^-} dxdy\right|, \tag{6}$$

而定理的假定就是(5)式，也就是

$$\iint_{D^+} ydxdy = -\iint_{D^-} ydxdy. \tag{7}$$

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

这就是说如果上半块对 x 轴的力矩等于下半块对 x 轴的力矩时，我们有不等式(6)．不等式(6)的 D^+，D^- 可颠倒，因而得到与定理等价的结论．

证明分三步进行：

1）对称化，也就是在保持面积和力矩不变的情况下，我们可以把 D 改变成为对称于 y 轴的形式，几何直观来说是：把每条平行于 x 轴的，两端在 D 的边界上的线段，水平移动，使其中点在 y 轴上，这样移动后所得出的 D' 对于 y 轴对称，而且面积与力矩都不变．

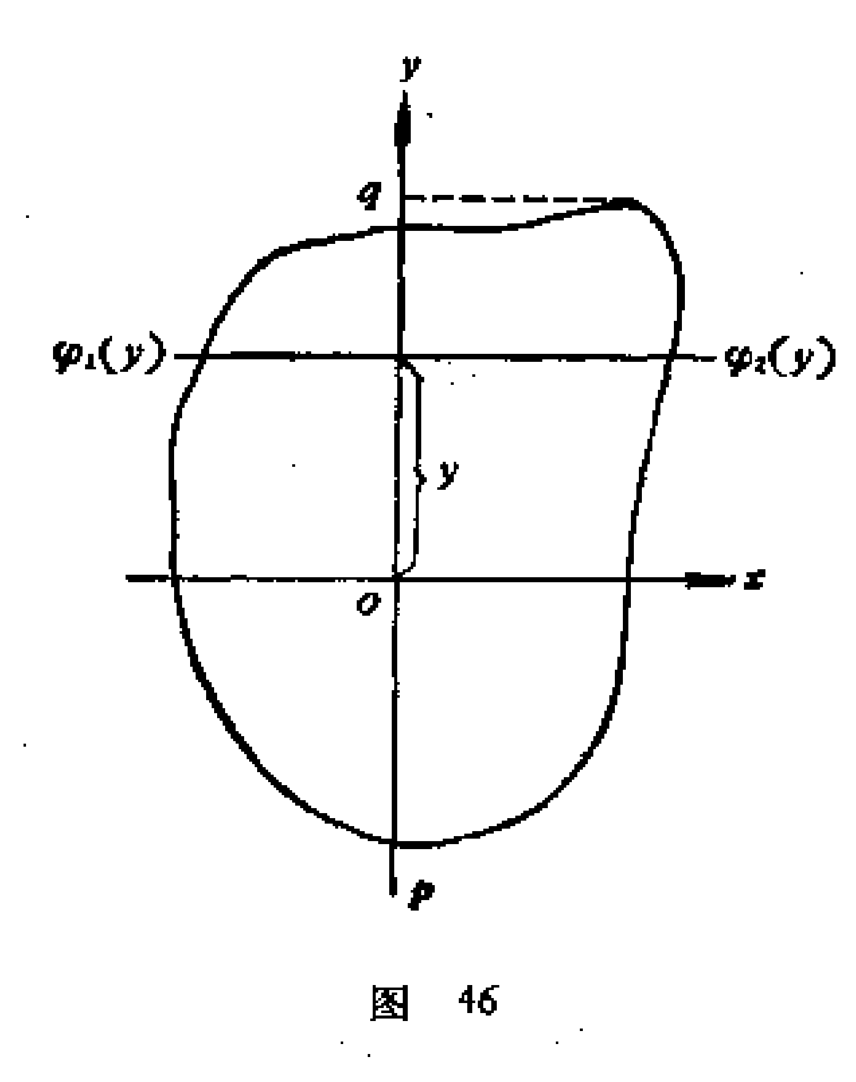


图 46

解析表示又怎样呢？我们看到，当 y 固定时，x 的积分范围是

$$\varphi_1(y) \leqslant x \leqslant \varphi_2(y),$$

因此

$$\iint_D dxdy = \int_p^q dy \int_{\varphi_1(y)}^{\varphi_2(y)} dx = \int_p^q (\varphi_2(y) - \varphi_1(y))dy$$

$$= \int_p^q dy \int_{-\frac{1}{2}(\varphi_2(y)-\varphi_1(y))}^{\frac{1}{2}(\varphi_2(y)-\varphi_1(y))} dx,$$

以及

$$\iint_D ydxdy = \int_p^q ydy \int_{-\frac{1}{2}(\varphi_2(y)-\varphi_1(y))}^{\frac{1}{2}(\varphi_2(y)-\varphi_1(y))} dx,$$

也就是我们的区域 D 变为

$$D': p \leqslant y \leqslant q, \quad -\frac{1}{2}(\varphi_2(y) - \varphi_1(y))$$

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

$$\leqslant x \leqslant \frac{1}{2}(\varphi_2(y) - \varphi_1(y)).$$

现在要证明 D' 也是凸的，也就是如果 (x_1, y_1), (x_2, y_2) 属于 D'，则

$$x = x_1 + t(x_2 - x_1), \quad y = y_1 + t(y_2 - y_1) \quad (0 \leqslant t \leqslant 1)$$

也属于 D'. 后一式是显然的. 由于

$$-\frac{1}{2}(\varphi_2(y_1) - \varphi_1(y_1)) \leqslant x_1 \leqslant \frac{1}{2}(\varphi_2(y_1) - \varphi_1(y_1)),$$

$$-\frac{1}{2}(\varphi_2(y_2) - \varphi_1(y_2)) \leqslant x_2 \leqslant \frac{1}{2}(\varphi_2(y_2) - \varphi_1(y_2)),$$

可知

$$-\frac{1}{2}[(1-t)(\varphi_2(y_1) - \varphi_1(y_1)) + t(\varphi_2(y_2) - \varphi_1(y_2))]$$

$$\leqslant x \leqslant \frac{1}{2}[(1-t)(\varphi_2(y_1) - \varphi_1(y_1)) + t(\varphi_2(y_2) - \varphi_1(y_2))].$$

因为 D 是凸的，所以

$$(1-t)\varphi_2(y_1) + t\varphi_2(y_2) \leqslant \varphi_2(y)$$
$$(1-t)\varphi_1(y_1) + t\varphi_1(y_2) \geqslant \varphi_1(y).$$

由此推出 (x, y) 也属于 D'.

因此，并不失去普遍性，我们可以假定 D 是对 y 轴对称的。

2) 拉直. x 轴交 D 的边界于 P, Q 二点，在 y 轴上找一点 R 作等腰三角形使其力矩等于 D^+ 的力矩. 我们证明三角形 PQR 的面积不大于 D^+ 的面积，也就是从

$$\iint_{D^+} y\,dx\,dy = \iint_{\triangle PQR} y\,dx\,dy$$

推出

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

$$\iint_{D^+} dxdy \geqslant \iint_{\triangle} dxdy.$$

三角形的高交 D^+ 的边界于 $y=y_0$，在三角形内而不在 D^+ 的点的 $y \geqslant y_0$，在 D^+ 不在三角形内的点的 $y \leqslant y_0$，即得所证.

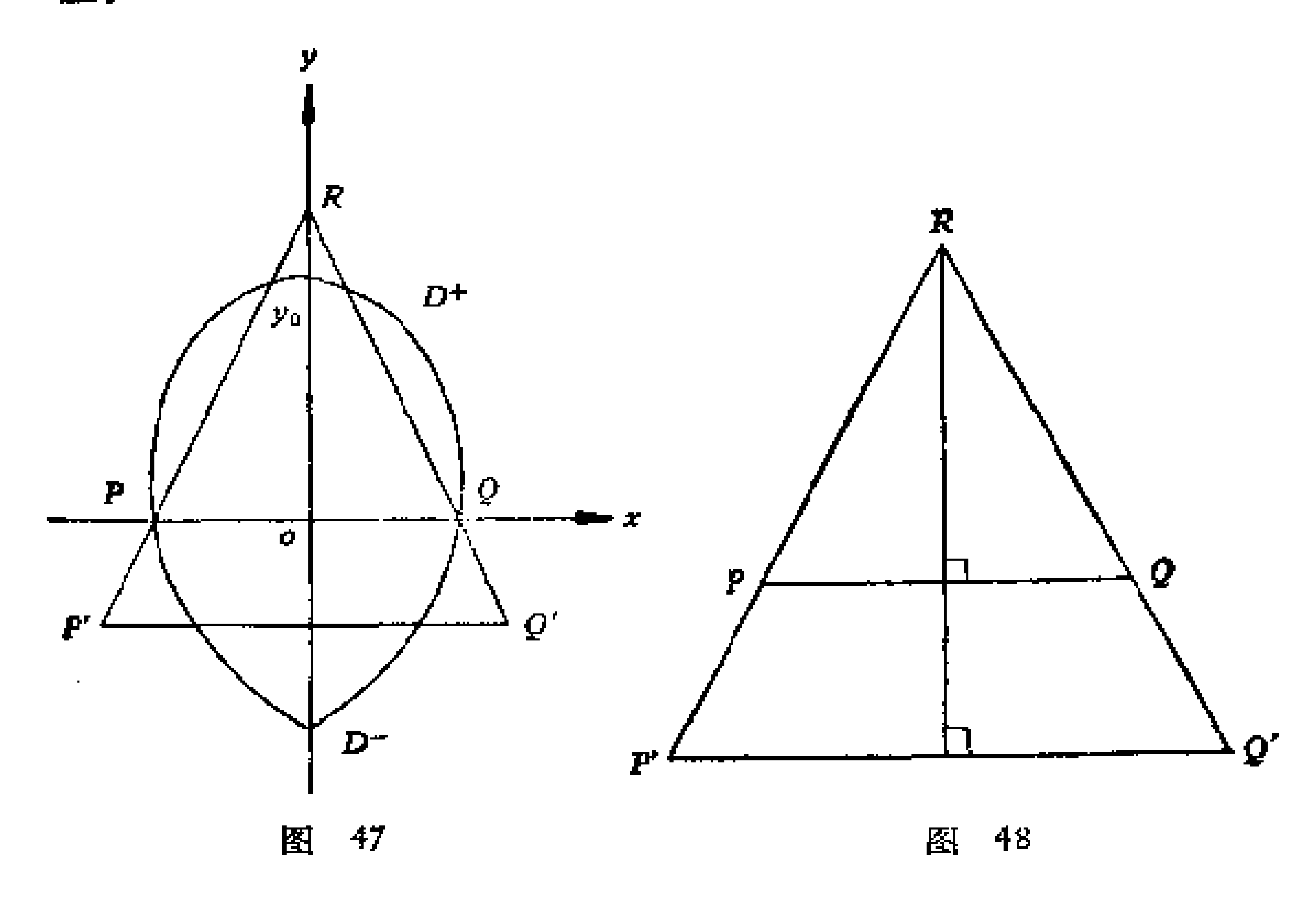


图 47　　　　图 48

再把 RP 和 RQ 分别延长到 P'，Q'，使得 $P'Q'$ 平行于 x 轴，使梯形 $PP'Q'Q$ 的力矩等于 D^- 的力矩，同样可推得

梯形 $PP'Q'Q$ 的面积 $\geqslant D^-$ 的面积，

因此

$$\left|\frac{D^+}{D^-}\right| \geqslant \frac{|\triangle PQR|}{|\text{梯形 } PP'Q'Q|},$$

由于力矩相等，$\triangle P'Q'R$ 的重心仍在 x 轴上.

因此，只需证明

3）等腰三角形过重心平行底边的直线分三角形为两部分

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

$$\frac{|\triangle PQR|}{|\text{梯形 } PP'Q'Q|}=\frac{4}{5}.$$

由于 R 到重心的距离等于高的 2/3，因此三角形 PQR 的面积是 $\triangle P'Q'R$ 的面积的 4/9，即得所证。

以下就依这三个步骤，把切块法推广到 s 个变数的情况。

§2. s 维锥形的体积与重心

s 维锥形的底盘是在 $x_s=0$ 的超平面上，半径为 r 的 $s-1$ 维球

$$x_1^2+\cdots+x_{s-1}^2\leqslant r^2,$$

其高为 h，当 $x_s=0$ 时，半径 ρ 等于 r，当 $x_s=h$ 时，半径 ρ 等于 0，边界的方程是

$$\rho=r\,\frac{h-x_s}{h}.$$

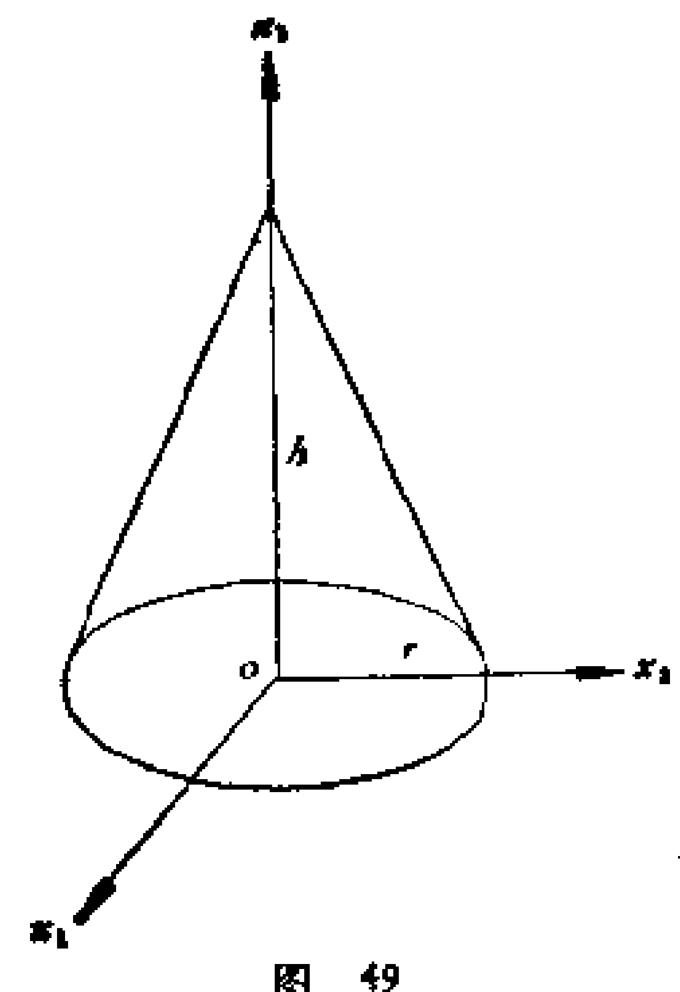


图 49

因此，锥形的体积

$$\begin{aligned}V_s&=\int_0^h dx_s\int\cdots\int_{x_1^2+\cdots+x_{s-1}^2\leqslant\rho^2} dx_1\cdots dx_{s-1}=\int_0^h\rho^{s-1}dx_s\int\cdots\int_{x_1^2+\cdots+x_{s-1}^2\leqslant 1} dx_1\cdots dx_{s-1}\\&=w_{s-1}\int_0^h\left(\frac{r}{h}(h-x_s)^{s-1}\right)dx_s=w_{s-1}r^{s-1}h\int_0^1(1-t)^{s-1}dt\\&=\frac{w_{s-1}r^{s-1}h}{s},\end{aligned}$$

这里 w_{s-1} 是 $s-1$ 维单位球的体积，这公式也可以述为：s

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

维锥的体积等于底积乘高的 s 分之一。

对超平面 $x_s=0$ 的力矩等于

$$\int\cdots\int x_s dx_1\cdots dx_s=\int_0^h x_s dx_s\int\cdots\int dx_1\cdots dx_{s-1}x_1^2$$

$$+\cdots+x_{s-1}^2\leqslant\rho^2=w_{s-1}r^{s-1}h^2\int_0^1(1-t)^{s-1}t\,dt$$

$$=w_{s-1}r^{s-1}h^2/s(s+1).$$

因此，重心的坐标是

$$\bar{x}_1=\bar{x}_2=\cdots=\bar{x}_{s-1}=0,\quad \bar{x}_s=\frac{h}{s+1}.$$

过重心平行于 $x_s=0$ 的超平面把 s 维锥分为两部分，其一是高为原高的 $1-\frac{1}{s+1}$ 倍，即依顶点缩小 $\frac{s}{s+1}$ 的圆锥，它是原来圆锥体积的

$$\left(\frac{s}{s+1}\right)^s$$

倍，因此圆锥被过重心的平面 $x_s=\frac{h}{s+1}$ 分为两部分，其体积之比是

$$\frac{\left(\frac{s}{s+1}\right)^s}{1-\left(\frac{s}{s+1}\right)^s},$$

这相当于上节的第三步。

§3. 对　称　化

我们假定 s 维凸体 D 被过重心的超平面

$$x_s=0$$

分为两部分，上、下半部各以 D^+，D^- 表之，则问题变为在

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

$$\int_{D^+}\cdots\int x_s dx_1\cdots dx_s = -\int_{D^-}\cdots\int x_s dx_1\cdots dx_s \tag{1}$$

的条件下，求比例

$$\int_{D^+}\cdots\int dx_1\cdots dx_s \Big/ \int_{D^-}\cdots\int dx_1\cdots dx_s \tag{2}$$

的上下界的问题。

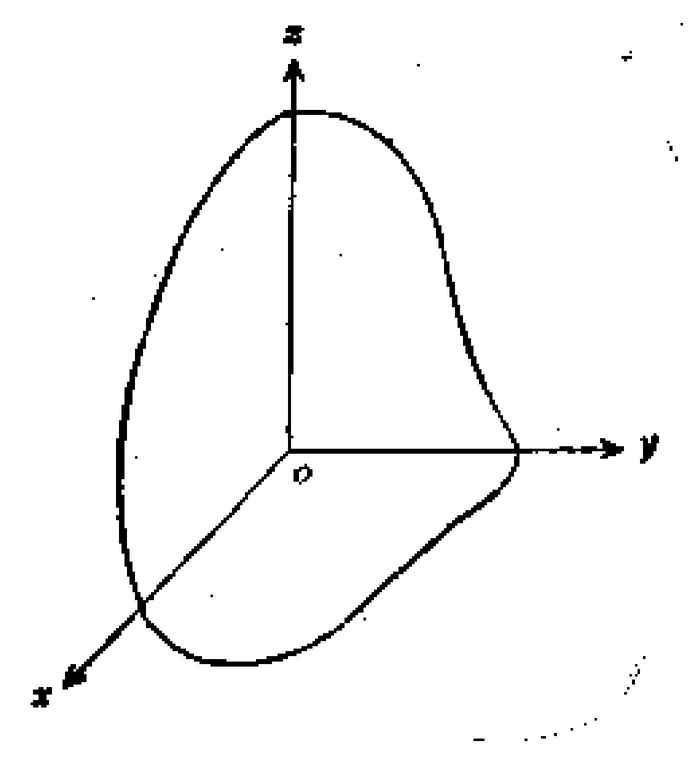


图 50

先用对称化的方法，不妨假定 D 是旋转体，再由“拉直”法来得出我们的结论。

为叙述简单，我们假定 $s=3$，将 (x_1, x_2, x_3) 写为 (x, y, z)。

依 xz 平面对称化，一垂直于 xz 平面的直线交凸体于二点 p, q，把这线段的中点移到 xz 平面上，则得出另一体 D_1，D_1 的 D_1^+，D_1^- 的体积和力矩仍然相等，即

$$\iiint_{D^+} z dz dx dy = \iiint_{D_1^+} z dz dx dy,$$

$$\iiint_{D^-} z dz dx dy = \iiint_{D_1^-} z dz dx dy,$$

$$\iiint_{D^+} dz dx dy = \iiint_{D_1^+} dz dx dy,$$

$$\iiint_{D^-} dz dx dy = \iiint_{D_1^-} dz dx dy. \tag{3}$$

这一点的证明是很容易的，因为当 x, z 定了，y 的积分范围是

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

$$\varphi(x,z)\leqslant y\leqslant \psi(x,z),$$

$$\iiint f(z)dzdxdy=\iint f(z)dzdx\int_{\varphi(x,z)}^{\psi(x,z)}dy.$$

因而可用类似于§2的方法证得(3)式，并且同样方法证得D_1也是凸的。

实际上，以上的方法可以用任何一个过z轴的平面来代替，这样我们得出一系列的凸域，这些凸域的极限是一个绕z轴的旋转体，为了证明这一点，我们讲得更具体些。

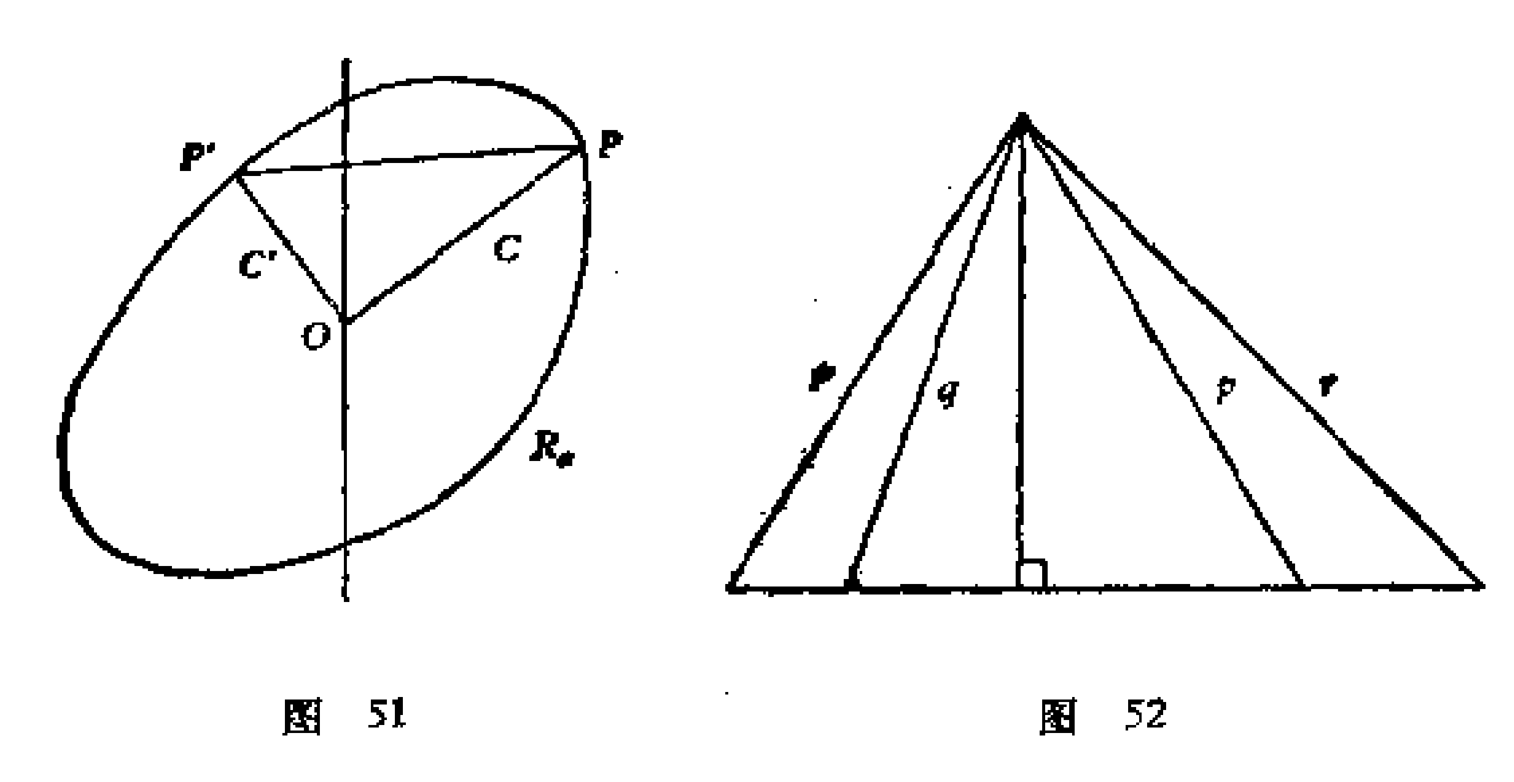


图 51　　　　图 52

先固定一个a，$z=a$截凸体得一平面凸域R_a，R_a的周界距原点最远、最近的距离分别为C与C'，点在P与P'。从原点作直线垂直于P'，P的联线，依这条直线对称化，得出来的图形长径一定比C小，短径一定比C'大(注意，由于等腰三角形顶点到底边上一点的联线一定比腰长短，到底边延线上一点的联线一定比腰长长，因此对称化总是把C缩小，把C'增大)。

由于得出一联串的C是递降的数列，因而有一极限C_0，一联串的C'是递增的数列，因而有一极限C_0'，如果$C_0\neq C_0'$，则以上的方法还是可以做下去，因此所得的极限是一个圆。

已经成为圆的对称化也还是它自己，再检查任意a，这做

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

法可使任意 $z=a$ 的截体都是圆.

但需要声明一下,一系列凸体的极限还是凸体,因而证明了:存在着一个绕 z 轴的旋转体,适合(3)式.

§4. Brun-Minkowski 不等式

对一凸旋转体来说: $z=a$ 所截得的圆的半径是 $\rho(a)$,面积是 $\pi\rho^2(a)=A(a)$,因此

$$\sqrt{A(a)}=\sqrt{\pi}\,\rho(a)$$

是一个凸函数,也就是有不等式

$$(1-\vartheta)\rho(a)+\vartheta\rho(b)\leqslant\rho((1-\vartheta)a+\vartheta b),\quad 0<\vartheta<1.$$

对 s 维来说,$x_s=a$ 所截下来的 $s-1$ 维体积以 $V(a)$ 表之,则有不等式

$$(1-\vartheta)V(a)^{\frac{1}{s-1}}+\vartheta V(b)^{\frac{1}{s-1}}$$
$$\leqslant V((1-\vartheta)a+\vartheta b)^{\frac{1}{s-1}}$$
$$(0<\vartheta<1).$$

这是有名的 Brun-Minkowski 不等式(见 Minkowski Geometrie der Zahlen, 数的几何).

另一个办法:先证明了这个不等式,然后以 a 为中心,作一体积为 $V(a)$ 的 $s-1$ 维球. 当 a 依 z 轴移动,则得出一个凸旋转体,而凸旋转体适合于上节公式(3)所示的关系.

§5. 拉　　直

主要定理　通过 s 维凸体 D 的重心的超平面把凸体分为两部分,其体积之比为 M,

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

$$M \leqslant \frac{1-\left(\frac{s}{s+1}\right)^{s}}{\left(\frac{s}{s+1}\right)^{s}}, \quad M \geqslant \frac{\left(\frac{s}{s+1}\right)^{s}}{1-\left(\frac{s}{s+1}\right)^{s}}.$$

事到如今，这个定理的证明就不难了．由前已知，只要假定这个 s 维凸体是重心在平行于底的超平面 $x_s=0$ 上的绕 x_s 轴的旋转体，作一个以 $x_s=0$ 交 D 的 $s-1$ 维球为底的 s 维锥体，它的高 h 使这个锥体的力矩等于 D^+ 的力矩．设这个锥体交 D 于 $x_s=l$，由于 D 外锥内的点适合于 $x_s>l$，而锥外 D^+ 内的点适合于 $x_s<l$，因而锥体的体积不大于 D^+ 的体积，把锥延展到下半空间，使“台”体的力矩等于 D^+ 的力矩，同法证明“台”体的体积不小于 D^- 的体积，因而

$$\frac{D^+\text{体积}}{D^-\text{体积}} \geqslant \frac{\text{锥体积}}{\text{“台”体积}} = \frac{\left(\frac{s}{s+1}\right)^{s}}{1-\left(\frac{s}{s+1}\right)^{s}}$$

即得所证．

§6. 说　　明

1）这个方法有很大的优越性，每次留下的部分不大于上次的

$$1-\left(\frac{s}{s+1}\right)^{s}$$

倍，由于

$$\left(\frac{s}{s+1}\right)^{s} > \frac{1}{e},$$

所以留下的不大于上次的

$$1-\frac{1}{e}$$

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

倍，这个数与 s 无关，而每次要做 $s+1$ 个试验，因而如果要求使留下的体积是原来的 $\frac{1}{m}$，那末试验次数的数量级是

$$s \log m,$$

即使用体积是长度的 s 方计算，也只要

$$s^2 \log m$$

数量级的次数计算就够了，这也就是一维的 s^2 倍.

2）这方法有两个缺点，一个是用差分

$$\frac{f(x_1+h, x_2, \cdots, x_n) - f(x_1, x_2, \cdots, x_n)}{h}$$

代替微商，“小数除小数”，这是计算数学中忌讳的，必须小心在意.

另一方面除 $s=2$ 外，重心的计算很复杂. 关于这一点可以求助于电子计算机，通过求积分的近代方法，问题可以化为球面上求积分来解决.

3）问题的实质是：凸体 D 中找一点 P，过这点作一 s 一

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

第四章　二次迴归法评介

述　　要

以下介绍一个较简单，较粗糙的方法，虽然如此，但比一般利用统计学上二次迴归求最大值的方法还可能好些。

方法是：作两条平行线 l_1，l_2，在这两条线上用单因素优选法各找出一取最大值的点 P_1，P_2，通过这两点作一直线 m，在 m 上取最大值的点 Q，就作为优选点了，所取的值就作为我们所要找的最大值。

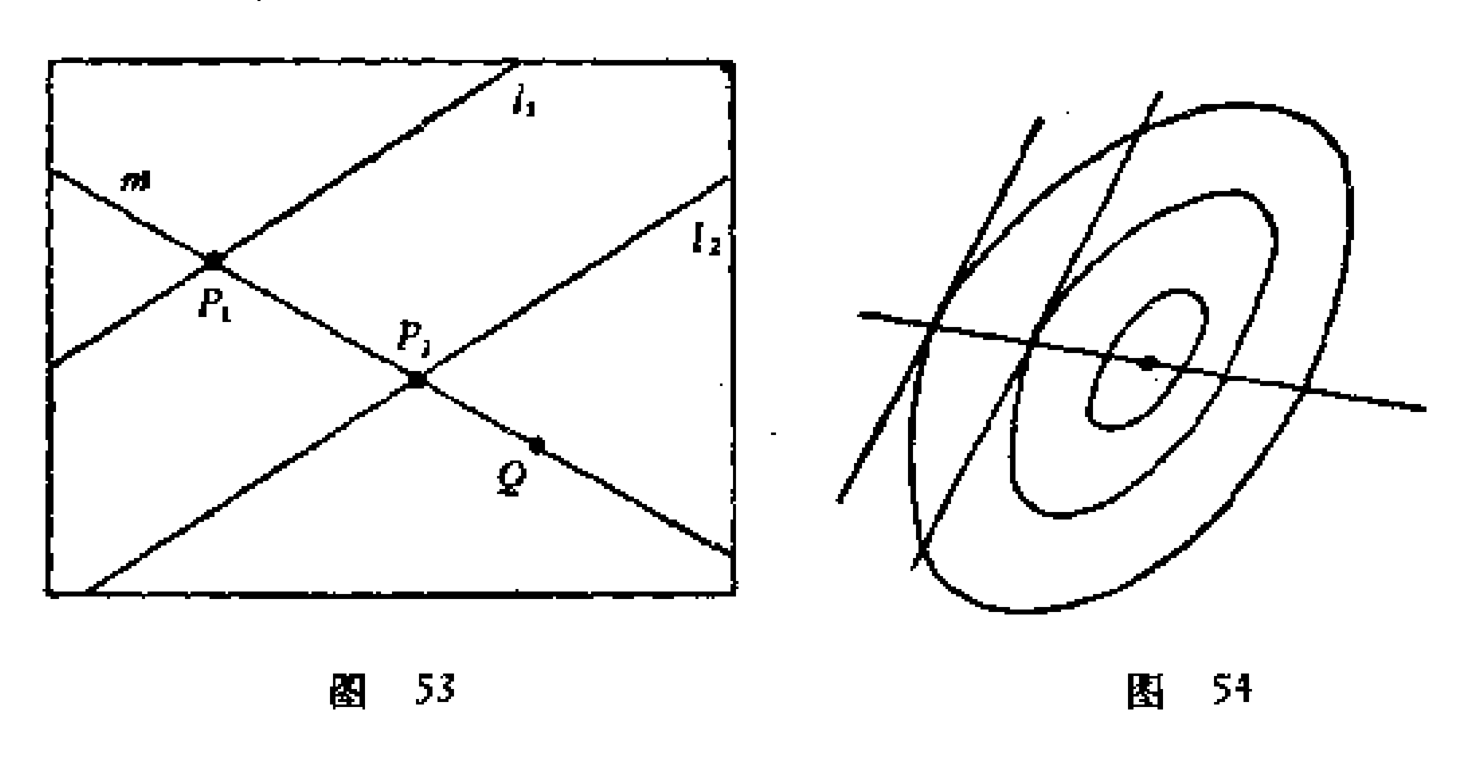


图　53　　　　　　图　54

这方法的局限性很大，但还是有些启发。

§1. 背　　景

正是从椭球的等高线图形——一簇以制高点为中心的同心椭圆启发出了这个方法。它的基础是：两个同心椭圆的两

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

条平行切线的切点联线一定通过中心。

其证明是：假定中心在原点，两个椭圆各为

$$ax^2 + 2bxy + cy^2 = \lambda,$$

$$ax^2 + 2bxy + cy^2 = \mu,$$

这里

$$ac - b^2 \neq 0.$$

他们在 (x_1, y_1), (x_2, y_2) 的切线各为

$$(ax_1 + by_1)x + (bx_1 + cy_1)y = \lambda,$$

$$(ax_2 + by_2)x + (bx_2 + cy_2)y = \mu.$$

如果他们平行，则

$$\frac{ax_1 + by_1}{ax_2 + by_2} = \frac{bx_1 + cy_1}{bx_2 + cy_2},$$

即

$$(ac - b^2)(x_1y_2 - x_2y_1) = 0.$$

由于 $ac - b^2 \neq 0$，所以

$$\frac{x_1}{y_1} = \frac{x_2}{y_2},$$

也就是 (x_1, y_1), (x_2, y_2), $(0, 0)$ 在一直线上.

本章开始所讲的方法，实质上是：P_1, P_2 是两个依中心放大的椭圆的切线的切点（并不排斥两个椭圆重合的情况），因而他们的联线通过中心。中心既是这联线上取最大值的点，也同时就是制高点了.

这些结果还可适当推广到稍为广一些的函数类. 假定

(1) $$f(x, y) \leqslant 1$$

是一凸域，并且适合

(2) $$f(xt, yt) = t^p f(x, y).$$

上面例中的同心椭圆族属于这类函数，又例如

$$f(x, y) = \sqrt[\alpha]{|x|^\alpha + |y|^\alpha} \quad (\alpha \geqslant 1).$$

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

满足上述条件〈1〉及〈2〉的 $z=f(x,y)$，等高线由

$$f(x,y)=\lambda$$

所形成，这是一批以原点为中心的凸曲线．从中心出发的某一方向

$$x=t\cos\theta,\quad y=t\sin\theta$$

上的一点 $t=t_0$ 处，曲线

$$f(x,y)=\lambda_0$$

$(\lambda_0=t_0^p f(\cos\theta,\ \sin\theta))$ 的切线斜率是

$$-\frac{\partial f}{\partial x}\Big/\frac{\partial f}{\partial y}=-\left(\frac{\partial f}{\partial t}\frac{\partial t}{\partial x}+\frac{\partial f}{\partial\theta}\frac{\partial\theta}{\partial x}\right)\Big/$$

$$\left(\frac{\partial f}{\partial t}\cdot\frac{\partial t}{\partial y}+\frac{\partial f}{\partial\theta}\frac{\partial\theta}{\partial y}\right)=-[p\cos\theta f(\cos\theta,\ \sin\theta)$$

$$-\sin\theta\frac{\partial}{\partial\theta}f(\cos\theta,\ \sin\theta)]\Big/[p\sin\theta f(\cos\theta,\ \sin\theta)$$

$$+\cos\theta\frac{\partial}{\partial\theta}f(\cos\theta,\ \sin\theta)].$$

这是与 t 无关的，也就是在这一方向的任何一点上等高线的切线都是平行的，这是上面讨论的一个推广．

§2. 二 次 迴 归

这个方法的限制性是极大的，但与统计学上二次迴归的方法比较，并不见得差．下面介绍一下统计学上的二次迴归．

我们预先假定所考虑的问题合乎二次模型，也就是假定"z 是 x, y 的二次式"，即

$$z-z^*=A(x-x^*)^2+2B(x-x^*)(y-y^*)$$
$$+C(y-y^*)^2 \tag{1}$$

x^*, y^*, z^*, A, B, C 是六个待定常数．如果我们在六个点上

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

作了试验，得出六组 x_i, y_i, $z_i(i=1, 2, \cdots, 6)$，六个待定常数就可以确定.

统计学者把试验中可能有误差的因素也考虑进去，他们试验 m 次，在 (x_i, y_i) 处试验出 z_i $(1\leqslant i\leqslant m)$ 来，然后求 A, B, C, x^*, y^*, z^*，使误差平方和

$$\Phi=\sum_{i=1}^{m}[z_i-z^*-A(x_i-x^*)^2-2B(x_i-x^*)(y_i-y^*)-C(y_i-y^*)^2]^2$$

最小.

由 (1) 可知，如果 $A<0$，且 $AC<B^2$，则显而易见在 $x=x^*$, $y=y^*$ 处，z 取最大值 z^*.

这个方法的计算比较麻烦，而且一开始就来了一个限制很强的数学模型 (1) 这一假定.

这模型的等高线是以 (x^*, y^*) 为相似中心的椭圆族，反之，等高线有相似椭圆族性质的曲面不一定是 (1) 的形式，因而，本章所研究的方法虽然特殊，但比二次迴归还较一般些.

预定数学模型是二次的，这的确是一个极为大胆的设想，它很少可能符合客观事实的，但是当接近最优解时，也就是 $x-x^*$, $y-y^*$ 充分小时，我们略去高次项，不妨就用 $x-x^*$, $y-y^*$ 的二次式来代替 z，而这正是难以辨别好坏的时候.

§3. s 个因素的问题

先讨论三个因素的情形:

开始取二个平行平面，在其中一个平面 π_1 用上面的方法，求三条直线上的最大值，得出在这平面上取最大值的点. 在另一个平面 π_2 就不需做三条直线上的优选法了，因为前一平面已定出了二个共轭方向 l_1 与 l_3. 我们在平面 π_2 上取一条平

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

行于 l_1 的直线 l'_1，在其上优选的最好点 P'_1，过 P'_1 作平行于 l_3 的直线 l'_3，在其上优选得最好点 P'_2，则 P'_2 便是平面 π_2 上的最好点了．再在连结 P'_2 与 π_1 上最好点 P_2 的直线上优选，便得出所求的最好点 P_0 来．

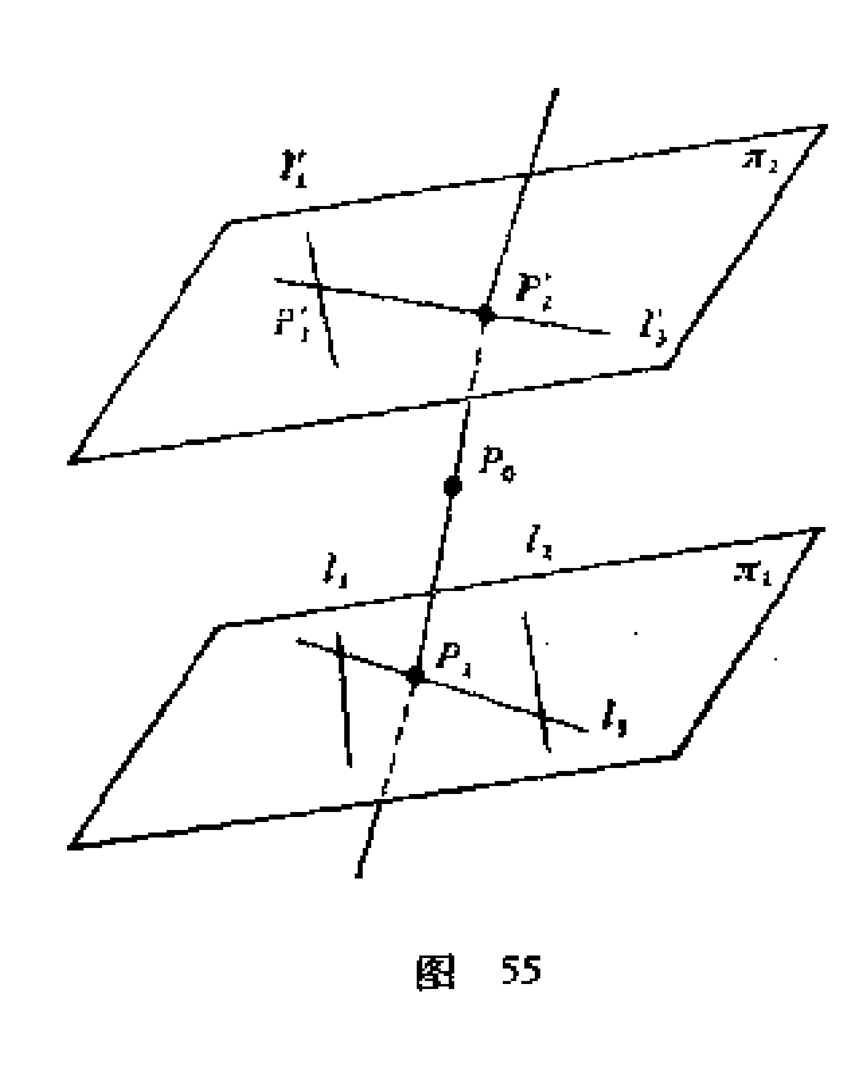


图 55

因而三个因素时，总共做了 $3+2+1=6$ 批单因素的优选试验，然后得出最大值．

我们很容易用归纳法证明，s 个因素需要做

$$\frac{s(s+1)}{2}$$

批单因素的试验．

如果单因素的试验用黄金分割法去做，次数的数量级仍然是 $\log\frac{1}{\delta}$，现在只是前面加上一个 $\frac{s(s+1)}{2}$ 的因子．

§4. 讨　　论

虽然屡次指出了这方法的局限性，但这个方法本身还是有所启发的．

首先，因子 2^s-1 实在太大，我们可以考虑以下的方法：

从一点 (x_1, y_1) 出发，联系切块法的思想，作一最陡上升的方向

$$\left(x_1+\frac{f(x_1+h, y_1)-f(x_1, y_1)}{h}t,\right.$$

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

$$y_1+\frac{f(x_1, y_1+k)-f(x_1, y_1)}{k}\Big).$$

垂直这个方向作一直线 l_2 在 l_2 上行单因素优选，得一点 (x_2, y_2)，再在 (x_1, y_1)，(x_2, y_2) 的联线上找最优点 R.

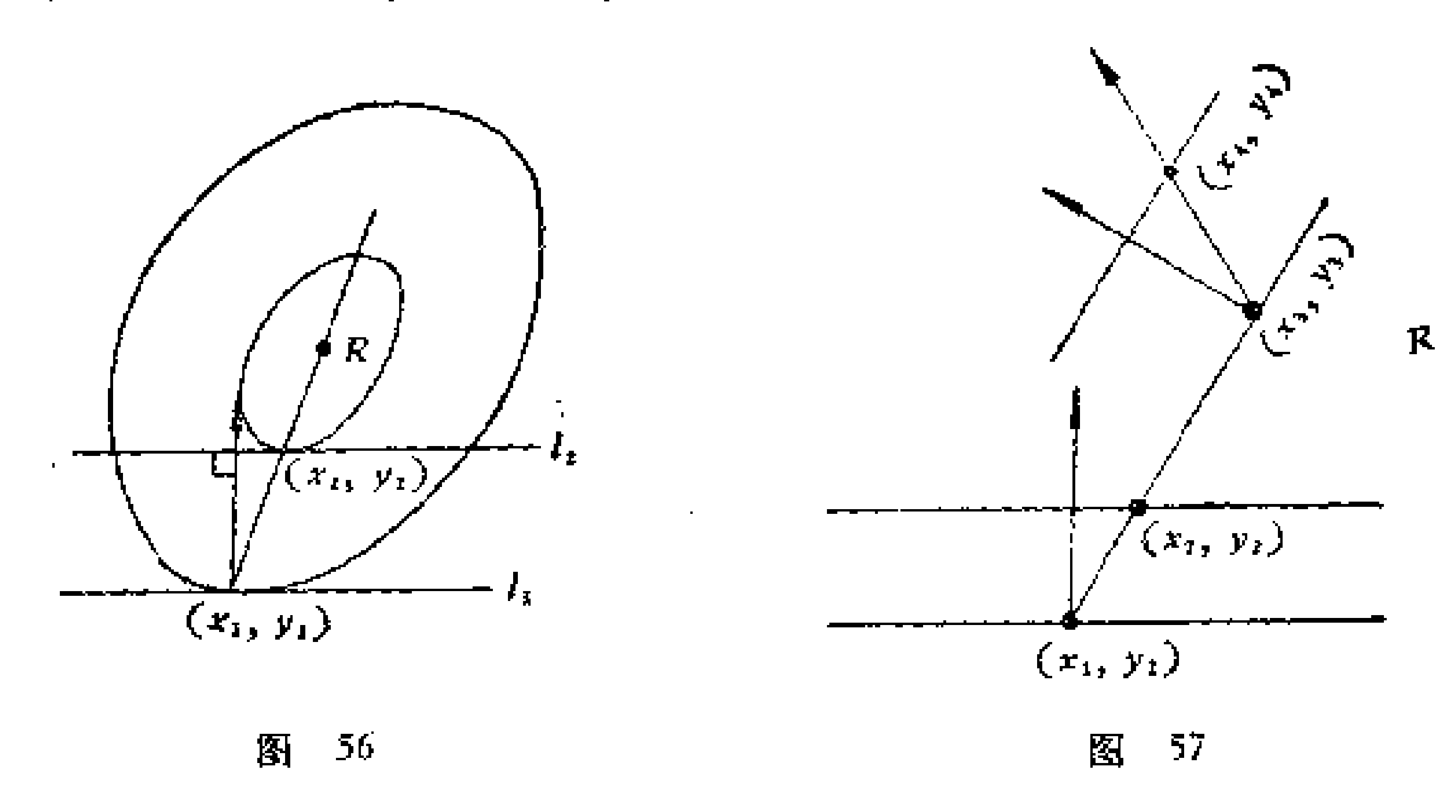


图 56　　　　图 57

这方法推广到 s 个因素，只要 s 个单因素优选法就行了.

其次，我们并不拿 R 作为最后数据，而作为再一次的开始，用这办法来逐步逼近最优解.

其三，改变迴旋上升法.

在一直线 l_1 上找到一最优点 P_1，通过 P_1 作垂线 l_2，在这垂线上用单因素法找到另一点 P_2，过 P_2 作 l_2 的垂线 l_3，在 l_3 上又找到点 P_3，接下去，不作 l_3 的垂线了，由于 P_1，P_3 是两等高线的平行切线的切点，因而联 P_1,P_3，在这直线 l_4 上找最优点 P_4，过 P_4 作 l_4 的垂线 l_5 等等.

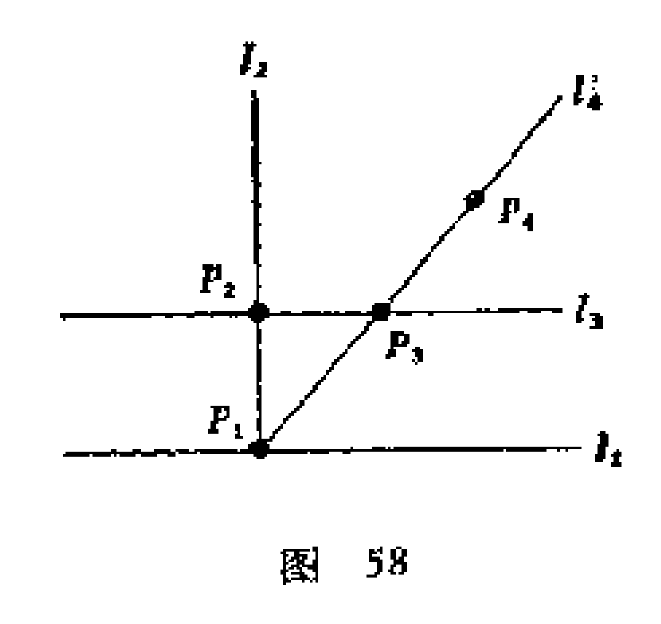


图 58

从多因素优选法可以看出，随着维数 s 的增加试验次数很快地增加，所以我们切不要因素愈多愈好而轻易地增加变

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

量，而应当抓住主要矛盾，先集中少数几个因素来做试验。这种突出主要矛盾的方法可以大大减少工作量。当主要矛盾解决之后，另一个矛盾上升为主要矛盾时，我们再用同样的方法处理。

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

第五章　抛物体法

述　　要

单因素抛物线法是利用三点的数据作出一抛物线，以达到此抛物线最大值的点作为新的试点，这方法所需次数的数量级是

$$\log\log\frac{1}{\varepsilon}.$$

把抛物线法的想法推广到多因素，便得出抛物体法.

以二个因素为例，根据在 $P_i(x_i, y_i)(i=1, 2, \cdots, 6)$ 点上的试验结果，可以确定一抛物面

$$z = F(x, y) = ax^2 + bxy + cy^2 + dx + ey + f.$$

求出此抛物面的顶点 (x^*, y^*, z^*)，然后在 (x^*, y^*) 作试验，与 P_2, P_3, P_4, P_5, P_6 的结果一起又可构成新的抛物面，…. 这样不断叠代.

我们将证明，当满足一定的条件时，这方法收敛速度是相当快的. 在 s 个因素的情形下所需次数的数量级仅为

$$s^2\log\log\frac{1}{\varepsilon}.$$

§1. 矩 阵 符 号

为了便于讲述本章的方法，我们引用矩阵符号，以 $A = A^{(l,m)}$ 表 l 行 m 列的矩阵：

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

$$A = A^{(l,m)} = (a_{ij}) \quad (1 \leqslant i \leqslant l, 1 \leqslant j \leqslant m). \qquad (1)$$

向量 $x = (x_1, x_2, \cdots, x_n)$ 可以看成为一行 n 列的矩阵 $x^{(1,n)}$. 我们常用 A' 表示由 A 经过行列互换所得的矩阵，它是 m 行 l 列的。

定义 我们称

$$\|A\| = \sum_{1\leqslant i\leqslant l}\sum_{1\leqslant j\leqslant m} |a_{ij}| \qquad (2)$$

为矩阵 A 的模。特别对于一个向量 x，我们用 $\|x\|$ 表示它的各个分量绝对值之和。这样定义的模显然有以下的性质:

$$\|A + B\| \leqslant \|A\| + \|B\|; \qquad (3)$$

$$\|AB\| \leqslant \|A\| \cdot \|B\|. \qquad (4)$$

当然，这些式子，只当 $A + B$, AB 有定义时才有意义.

引理 1 命 $X = X^{(n,s)}$, $A = A^{(s,s)}$, 若

$$\|A\| \leqslant \alpha < 1, \qquad (5)$$

则

$$\|X(I - A)^{-1}\| \leqslant 1/(1-\alpha)\|X\|. \qquad (6)$$

证明 我们有

$$X(I-A)^{-1} = X(I + A + A^2 + \cdots),$$

由 (3) 及 (4) 可知

$$\|X(I-A)^{-1}\| \leqslant \|X\| + \|XA\| + \|XA^2\| + \cdots$$
$$\leqslant \|X\|(1 + \alpha + \alpha^2 + \cdots) = 1/(1-\alpha)\|X\|.$$

§2. 方法的背景

我们的问题是：求 s 个变数 $x = (x_1, x_2, \cdots, x_s)$ 的函数 $\varphi(x)$ 的最大值.

在 $\frac{1}{2}(s+1)(s+2)$ 个点

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

$$x^{(0)}, \tag{1}$$

$$x^{(0)} + \eta e_i \quad (1 \leqslant i \leqslant s),$$

$$x^{(0)} + \eta(e_i + e_j) \quad (1 \leqslant i \leqslant j \leqslant s)$$

上各做一次试验，此处

$$e_i = (\underbrace{0, \cdots, 0}_{i-1\text{个}}, 1, 0, \cdots, 0),$$

得出 $\frac{1}{2}(s+1)(s+2)$ 个数据：

$$\varphi(x^{(0)}), \quad \varphi(x^{(0)} + \eta e_i), \quad \varphi(x^{(0)} + \eta(e_i + e_j)). \tag{2}$$

我们现在找一个二次式

$$\psi(x) = a + 2bx' + xCx', \tag{3}$$

其中 $a = a^{(1,1)}$, $b = b^{(1,s)}$, $C = C' = C^{(s,s)}$ 为一定负矩阵，使在 (1) 中各点上有

$$\psi(x) = \varphi(x) \tag{4}$$

成立，现在总共有 $\frac{1}{2}(s+1)(s+2)$ 个方程，由此可以定出 $\frac{1}{2}(s+1)(s+2)$ 个未知量 a, b, C.

令

$$\Delta_i f(x) = \frac{1}{\eta}[f(x + \eta e_i) - f(x)], \tag{5}$$

则

$$\begin{aligned}
\Delta_i \varphi(x^{(0)}) = \Delta_i \psi(x^{(0)}) &= \frac{1}{\eta}[2b(x^{(0)} + \eta e_i)' - 2b(x^{(0)})' \\
&\quad + (x^{(0)} + \eta e_i)C(x^{(0)} + \eta e_i)' - x^{(0)}Cx^{(0)\prime}] \\
&= 2be_i' + 2x^{(0)}Ce_i' + \eta e_i C e_i',
\end{aligned} \tag{6}$$

及

$$\Delta_i \Delta_j \varphi(x^{(0)}) = 2e_i C e_j'. \tag{7}$$

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

由(7)得出

$$C = \frac{1}{2}(\Delta_i \Delta_j \varphi(x^{(0)}))_{1 \leqslant i, j \leqslant s}. \tag{8}$$

由(6)得出

$$(\Delta_1 \varphi(x^{(0)}), \cdots, \Delta_s \varphi(x^{(0)})) = 2b + 2x^{(0)}C + \frac{1}{2}\eta(\Delta_1^2 \varphi(x^{(0)}), \cdots, \Delta_s^2 \varphi(x^{(0)}))$$

即得

$$2b = (\Delta_1 \varphi(x^{(0)}) - \frac{1}{2}\eta \Delta_1^2 \varphi(x^{(0)}), \cdots, \Delta_s \varphi(x^{(0)}) - \frac{1}{2}\eta \Delta_s^2 \varphi(x^{(0)})) - 2x^{(0)}C. \tag{9}$$

再由于 C 是定负的,及

$$\psi(x) = \alpha - bC^{-1}b' + (x + bC^{-1})C(x + bC^{-1})',$$

所以当

$$x = x^{(1)} = -bC^{-1}$$

时,$\psi(x)$ 取最大值,因此

$$x^{(1)} = x^{(0)} - V(x^{(0)})M(x^{(0)})^{-1}, \tag{10}$$

这里

$$V(x^{(0)}) = (\Delta_1 \varphi(x^{(0)}) - \frac{1}{2}\eta \Delta_1^2 \varphi(x^{(0)}), \cdots, \Delta_s \varphi(x^{(0)}) - \frac{1}{2}\eta \Delta_s^2 \varphi(x^{(0)})),$$

$$M(x^{(0)}) = (\Delta_i \Delta_j \varphi(x^{(0)}))_{1 \leqslant i, j \leqslant s}.$$

我们的方法是利用递归公式

$$x^{(\nu+1)} = x^{(\nu)} - V(x^{(\nu)})M(x^{(\nu)})^{-1} \tag{11}$$

来逐步逼近 $\varphi(x)$ 的最大值,此处

$$V(x^{(\nu)}) = (\Delta_1 \varphi(x^{(\nu)}) - \frac{1}{2}\eta \Delta_1^2 \varphi(x^{(\nu)}), \cdots, \Delta_s \varphi(x^{(\nu)}) - \frac{1}{2}\eta \Delta_s^2 \varphi(x^{(\nu)})), \tag{12}$$

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

及

$$M(x^{(\nu)}) = (\Delta_i\Delta_j\varphi(x^{(\nu)}))_{1\leqslant i,\ j\leqslant s}. \tag{13}$$

假定这最大值在 $x = x^*$ 处取，则

$$\frac{\partial}{\partial x_i}\varphi(x)|_{x=x^*} = 0 \quad (i = 1, 2, \cdots, s). \tag{14}$$

为简单起见，以后把 $\frac{\partial}{\partial x_i}\varphi(x)$ 记作 φ_{x_i}.

在证明这方法的有效性时，我们要用到下节两条定理.

§3. 两条定理

定理 1 如果

$$\sum_{i=1}^{s}|\varphi_{x_i x_j x_k}| \leqslant K, \quad \sum_{i=1}^{s}|\varphi_{x_i^2 x_k}| \leqslant K, \tag{15}$$

则

$$\|x^{(\nu+1)} - x^{(\nu)}\| \leqslant \frac{1}{2}K(\|x^{(\nu-1)} - x^{(\nu)}\|^2$$
$$+ 2\eta\|x^{(\nu-1)} - x^{(\nu)}\|)\|M(x^{(\nu)})^{-1}\|. \tag{16}$$

证 由(11)式得

$$(x^{(\nu+1)} - x^{(\nu)})M(x^{(\nu)}) = -V(x^{(\nu)}),$$
$$(x^{(\nu)} - x^{(\nu-1)})M(x^{(\nu-1)}) = -V(x^{(\nu-1)}),$$

因此得

$$(x^{(\nu+1)} - x^{(\nu)})M(x^{(\nu)}) = V(x^{(\nu-1)}) - V(x^{(\nu)})$$
$$+ (x^{(\nu)} - x^{(\nu-1)})M(x^{(\nu-1)}). \tag{17}$$

注意到

$$f(x^{(\mu)}) - f(x^{(\nu)}) = \int_0^1 \frac{d}{d\alpha}f(x^{(\nu)} + \alpha(x^{(\mu)} - x^{(\nu)}))d\alpha$$

$$= \sum_{k=1}^{s}(x_k^{(\mu)} - x_k^{(\nu)})\int_0^1 f_{x_k}(x^{(\nu)} + \alpha(x^{(\mu)} - x^{(\nu)}))d\alpha, \tag{18}$$

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

及

$$\Delta_i f(x^{(\nu)}) = \int_0^1 f_{x_i}(x^{(\nu)} + \alpha e_i \eta) d\alpha \tag{19}$$

(17)式右边矢量的第 i 个分量是

$$\left(\Delta_i - \frac{1}{2}\eta\Delta_i^2\right)\Big(\varphi(x^{(\nu-1)}) - \varphi(x^{(\nu)})$$

$$+ \sum_{k=1}^{s} (x_k^{(\nu)} - x_k^{(\nu-1)})\Delta_k\Delta_i\varphi(x^{(\nu-)})\Big)$$

$$= \sum_{k=1}^{s} (x_k^{(\nu-1)} - x_k^{?(\nu)}) \Big\{\left(\Delta_i - \frac{1}{2}\eta\Delta_i^2\right)\int_0^1 \varphi_{x_k}(x^{(\nu)}$$

$$+ \alpha(x^{(\nu-1)} - x^{(\nu)}))d\alpha$$

$$- \Delta_i \int_0^1 \varphi_{x_k}(x^{(\nu-1)} + \alpha\eta e_k) d\alpha\Big\}$$

$$= \sum_{k=1}^{s} (x_k^{(\nu-1)} - x_k^{(\nu)}) \Big\{\int_0^1\int_0^1 \varphi_{x_k x_i}(x^{(\nu)}$$

$$+ \alpha(x^{(\nu-1)} - x^{(\nu)}) + \beta\eta e_i)d\alpha d\beta$$

$$- \frac{1}{2}\eta \int_0^1\int_0^1\int_0^1 \varphi_{x_i^2 x_k}(x^{(\nu)} + \alpha(x^{(\nu-1)} - x^{(\nu)})$$

$$+ \beta\eta e_i + \gamma\eta e_i)d\alpha d\beta d\gamma$$

$$- \int_0^1\int_0^1 \varphi_{x_k x_i}(x^{(\nu-1)} + 2\eta e_k + \beta\eta e_i)d\alpha d\beta\Big\}$$

$$= \sum_{k=1}^{s} (x_k^{(\nu-1)} - x_k^{(\nu)})\Big\{\sum_{l=1}^{s} \int_0^1\int_0^1\int_0^1 [(1-\alpha)(x_l^{(\nu)}$$

$$- x_l^{(\nu-1)}) - \alpha\eta\delta_{kl}]\varphi_{x_k x_l x_i}[x^{(\nu-1)}$$

$$+ \alpha\eta e_k + \beta\eta e_i + \gamma((1-\alpha)(x^{(\nu)} - x^{(\nu-1)})$$

$$- \alpha\eta e_k]\, d\alpha d\beta d\gamma - \frac{1}{2}\eta \int_0^1\int_0^1\int_0^1 \varphi_{x_i^2 x_k}(x^{(\nu)}$$

$$+ \alpha(x^{(\nu-1)} - x^{(\nu)}) + \beta\eta e_i + \gamma\eta e_i)d\alpha d\beta d\gamma\Big\},$$

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

其中 $\delta_{kl}=\begin{cases}0 & k\neq l,\\ 1 & k=l.\end{cases}$ 由假定(15)，可知

$$\|x^{(\nu+1)}-x^{(\nu)}\|\leqslant\|x^{(\nu+1)}-x^{(\nu)}\|\cdot\Big\{K\|x^{(\nu)}-x^{(\nu-1)}\|$$

$$\cdot\int_0^1\int_0^1\int_0^1(1-\alpha)d\alpha d\beta d\gamma+\eta\bar{K}\int_0^1\int_0^1\int_0^1\alpha d\alpha d\beta d\gamma$$

$$+\frac{1}{2}\eta K\int_0^1\int_0^1\int_0^1 d\alpha d\beta d\gamma\Big\}\|M(x^{(\nu)})^{-1}\|$$

$$=\frac{K}{2}(\|x^{(\nu-1)}-x^{(\nu)}\|^2$$

$$+2\eta\|x^{(\nu-1)}-x^{(\nu)}\|)\|M(x^{(\nu)})^{-1}\|.$$

这就是(16)式。

定理 2 如果

$$\sum_{i=1}^{s}|\varphi_{x_ix_jx_k}|\leqslant K,$$

及

$$K\|x^{(\nu)}-x^{(\nu-1)}\|\cdot\|M(x^{(\nu-1)})^{-1}\|\leqslant\alpha_{\nu-1}<1,\qquad(20)$$

则

$$\|M(x^{(\nu)})^{-1}\|\leqslant\frac{1}{1-\alpha_{\nu-1}}\|M(x^{(\nu-1)})^{-1}\|.\qquad(21)$$

证 命

$$M(x^{(\nu)})-M(x^{(\nu-1)})=(p_{ij}),$$

$$M(x^{(\nu-1)})^{-1}=(q_{ij}),$$

则

$$p_{ij}=\Delta_i\Delta_j(\varphi(x^{(\nu)})-\varphi(x^{(\nu-1)}))$$

$$=\Delta_i\Delta_j\int_0^1\frac{d}{d\alpha}\varphi[x^{(\nu-1)}+\alpha(x^{(\nu)}-x^{(\nu-1)})]d\alpha$$

$$=\sum_{k=1}^{s}(x_k^{(\nu)}-x_k^{(\nu-1)})\Delta_i\Delta_j$$

$$\cdot\int_0^1\varphi_{x_k}(x^{(\nu-1)}+\alpha(x^{(\nu)}-x^{(\nu-1)}))d\alpha,$$

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

因此

$$\|I - M(x^{(\nu)})M(x^{(\nu-1)})^{-1}\| \leqslant \sum_{i=1}^{s}\sum_{j=1}^{s}\sum_{k=1}^{s}|p_{ij}|\cdot|q_{jk}|$$

$$\leqslant \sum_{i=1}^{s}\sum_{j=1}^{s}\sum_{k=1}^{s}\sum_{l=1}^{s}|x_k^{(\nu)} - x_k^{(\nu-1)}|$$

$$\cdot\int_0^1 |\Delta_i\Delta_j\varphi_{x_k}(x^{(\nu-1)} + \alpha(x^{(\nu)} - x^{(\nu-1)}))|d\alpha\cdot|q_{jl}|$$

$$\leqslant K\|x^{(\nu)} - x^{(\nu-1)}\|\cdot\|M(x^{(\nu-1)})^{-1}\|. \tag{22}$$

命

$$I - M(x^{(\nu)})M(x^{(\nu-1)})^{-1} = A. \tag{23}$$

则

$$M(x^{(\nu)})^{-1} = M(x^{(\nu-1)})^{-1}(I - A)^{-1},$$

由假定(20)及引理1得出本定理.

§4. 有 效 性

定理 3 如果

$$K\|x^{(1)} - x^{(0)}\|\|M(x^{(0)})^{-1}\| \leqslant \alpha_0 < \frac{3-\sqrt{5}}{2}$$

$$(=0.382\cdots), \tag{24}$$

$$\|x^{(\nu)} - x^{(\nu-1)}\| \geqslant 2\eta, \tag{25}$$

则

$$\|x^{(\nu+1)} - x^{(\nu)}\|$$

$$\leqslant \left(\frac{3+\sqrt{5}}{2}\alpha_0\right)^{2\nu-1}\left(\frac{\sqrt{5}-1}{2}\right)^{\nu}\|x^{(1)} - x^{(0)}\|. \tag{26}$$

证 1) 由定理1, 2及(25)可知

$$K\|x^{(\nu+1)} - x^{(\nu)}\|\|M(x^{(\nu)})^{-1}\|$$

$$\leqslant (K\|x^{(\nu)} - x^{(\nu-1)}\|\|M(x^{(\nu)})^{-1}\|)^2$$

$$\leqslant \left(\frac{\alpha_{\nu-1}}{1-\alpha_{\nu-1}}\right)^2 = \alpha_\nu.$$

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

现在用归纳法证明

$$0 < \alpha_\nu < \frac{3-\sqrt{5}}{2}. \tag{27}$$

由于$f(x) = \left(\frac{x}{1-x}\right)^2$是$x$的递增函数$(0 < x < 1)$，所以对于不大于$\frac{3-\sqrt{5}}{2}$的$x$可有

$$f(x) \leqslant f\left(\frac{3-\sqrt{5}}{2}\right) = \left(\frac{3-\sqrt{5}}{2} \Big/ \left(\frac{-1+\sqrt{5}}{2}\right)\right)^2$$

$$= \left(\frac{3-\sqrt{5}}{2}\right)^2 \Big/ \left(\frac{3-\sqrt{5}}{2}\right) = \frac{3-\sqrt{5}}{2}.$$

于是(27)得证.

2）由定理2，可知

$$\|M(x^{(\nu)})^{-1}\| \leqslant \frac{1}{1-\alpha_{\nu-1}}\|M(x^{(\nu-1)})^{-1}\|$$

$$\leqslant \left(\frac{\sqrt{5}+1}{2}\right)\|M(x^{(\nu-1)})^{-1}\|$$

$$\leqslant \left(\frac{\sqrt{5}+1}{2}\right)^\nu \|M(x^{(0)})^{-1}\|. \tag{28}$$

再用归纳法证明(26)式．当$\nu = 0$，此式显然成立．由定理1，

$$\|x^{(\nu+1)} - x^{(\nu)}\| \leqslant K\|x^{(\nu)} - x^{(\nu-1)}\|^2\|M(x^{(\nu)})^{-1}\|$$

$$\leqslant K\left[\left(\frac{3+\sqrt{5}}{2}\alpha_0\right)^{2^{\nu-1}-1} - 1\left(\frac{\sqrt{5}-1}{2}\right)^{\nu-1}\right.$$

$$\left.\cdot \|x^{(1)} - x^{(0)}\|\right]^2 \cdot \left(\frac{\sqrt{5}+1}{2}\right)^\nu \|M(x^{(0)})^{-1}\|$$

$$\leqslant \left(\frac{3+\sqrt{5}}{2}\alpha_0\right)^{2^\nu-2}\left(\frac{\sqrt{5}-1}{2}\right)^{2\nu-2}$$

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

$$\cdot \|x^{(1)} - x^{(0)}\| \left(\frac{\sqrt{5}+1}{2}\right)^{\nu} \alpha_0.$$

由于

$$\left(\frac{3+\sqrt{5}}{2}\right)^{-1}\left(\frac{\sqrt{5}-1}{2}\right)^{-2} = 1$$

即得 (26) 式。

定理 3 说明，一旦找到一个初始值适合于条件 (24) 之后，我们这个方法进行得是非常迅速的，其收敛因子是

$$\left(\frac{1}{2}(3+\sqrt{5})\alpha_0\right)^{2^{\nu}-1}\left(\frac{1}{2}(\sqrt{5}-1)\right)^{\nu}.$$

第一个是主要的，第二个因子和黄金分割法差不多好，而且维数增加，试验次数只增加 s^2 倍，而不是 s 次方，如果取

$$\alpha_0 = \sqrt{5} - 2 \text{ 则 } \left(\frac{3+\sqrt{5}}{2}\alpha_0\right)^{2^{\nu}-1} = \left(\frac{\sqrt{5}-1}{2}\right)^{2^{\nu}-1}.$$

当

$$\nu = \left(\log\log\frac{1}{\varepsilon} - \log\log\frac{1+\sqrt{5}}{2}\right)\Big/\log 2$$

时，就可以使

$$\|x^{(\nu+1)} - x^{(\nu)}\| \leqslant \varepsilon\|x^{(1)} - x^{(0)}\|.$$

注意：在

$$\|x^{(\nu)} - x^{(\nu-1)}\| < 2\eta$$

时，我们停止做下去，此时，我们已经尽可能地达到 x^* 的渐近值了。

附记 也许有人提出条件 (24) 不易验证，我们的建议先用其他方法做到相当程度后，再用此法，在用此法时，如果试验出的 $\varphi(x^{(1)})$ 比 $\varphi(x^{(0)})$ 大，则照样做下去。当 $\varphi(x^{(1)})$ 不比 $\varphi(x^{(0)})$ 大时，可以与切块法同时考虑，因为在 $x^{(0)}$，$x^{(0)}+\eta l_i$，$x^{0}+\eta(l_i+l_j)$ 做了实验，可以来检查，$x^{(0)}$，$x^{(0)}+\eta l_i$（($s+1$) 个点）的上升方向，而另外 ($s+1$) 个点，$x^{(1)}$，$x^{(0)}+\eta l_i$ 又有

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

了 $s+1$ 个新方向，根据这些线索可以得出较小的区域，但有一点理论上是肯定的，充分多次之后，这方法便一帆风顺了.

§5. 补 充 方 法

在多变数时，以下的简单方法也能起到好的作用. 我们用叠代公式

(1) $x^{(\nu+1)}=x^{(\nu)}-u(x^{(\nu)})M(x^{(0)})^{-1}, \quad (\nu=1,2,\cdots)$

也就是 $M(x^{(0)})^{-1}$ 并不变化，$V(x)$ 代以更简单的

$$u=(u_1, u_2, \cdots, u_s),$$

$$u_i=u_i(x)=\Delta_i\varphi(x), \quad (i=1, 2, \cdots, s).$$

类似于前述，容易证明这时有

$$\|x^{(\nu+1)}-x^{(\nu)}\| \leqslant K\|x^{(\nu)}-x^{(\nu-1)}\|\Big(\|x^{(\nu)}-x^{(0)}\|$$

$$+\frac{1}{2}\|2^{(\nu)}-2^{(\nu-1)}\|+\frac{\eta}{2}\Big)\cdot\|M(x_0)^{-1}\|.$$

为了便于比较，我们略去了 η，又记 x^* 为制高点.

概括地说，抛物体法每次要做

$$\frac{1}{2}(s+1)(s+2)$$

个实验来定出

$$\varphi(x_0),\ \varphi(x_0+\eta l_i),\ \varphi(x_0+\eta l_i+\eta l_j)(1\leqslant i\leqslant j\leqslant s)$$

的数据，而收敛情形是二次的，也不难证明

$$\|x^{(\nu+1)}-x^*\| \ll \|x^{(\nu)}-x^*\|^2,$$

这就是说，收敛指数成倍地增长着.

补充方法是每次做 $s+1$ 个实验来定出

$$\varphi(x), \quad \varphi(x+\eta l_i) \quad (1\leqslant i\leqslant s)$$

的数据，但收敛情形是

$$\|x^{(\nu+1)}-x^2\| \ll \|x^{(\nu)}-x^*\|\|x^{(0)}-x^*\|$$

即收敛指数上升一阶.

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

第六章　与计算数学的关系

§1. 问题的叙述

如果 $f(x, y)$ 是一个有了表达式的函数，则求

$$z = f(x, y)$$

的最大值的计算问题，和我们所讨论的问题一样，但"试验"已经可以更具体地理解为"给了 x, y 的数值，通过计算得出 z 的数值"。由于近代电子计算机速度很快，因而多做些"试验"是完全可能的。如果计算一个 z 的数值需要一千次，在每秒十万次的计算机上只不过 1/100 秒的时间而已。但如果我们的 x, y 各要求算四位，且对每一对 (x, y) 算一个 z，需运算千次，然后比较出最大值，这样总共的计算次数就是

$$10^4 \times 10^4 \times 10^3 = 10^{11},$$

即一千亿次，这样每秒十万次的电子计算机还要计算三百个工作小时，太多了！

如果变数多了，增长率更快，电子计算机虽快，但还是不能胜任。

工作可考虑分两方面来进行，一方面搞电子计算机的往加快速度方面努力，使每秒钟的运算次数愈多愈好，愈可靠愈好，另一方面搞计算的，从减少计算次数方面努力，想方设法来减少计算次数。

前面各章所介绍的优选法，实质上也就是求最大值（或最小值）的快速计算法。实际上不仅如此，很多计算问题，都可以化为极值问题来解决，而极值问题离散化后，就与本书所讲

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

的方法有关了，因而优选法在计算数学上也有其重要性。

例 解超越联立方程

$$f(x, y)=0, \quad g(x, y)=0,$$

可以化为求函数

$$F(x, y)=f^2+g^2$$

的最小值的问题。

所以，我们不能把优选法简单地理解为一个减少试验次数的方法。实际上，这个方法给计算数学也增添了内容。

§2. 黄金分割法的计算格式

求方程，

$$y=f(x)=0$$

的解。假定在 (a, b) 中仅有一解（或本法仅能找出其诸解之一），计算步骤如下：

1）写下 $a, b, w=\frac{\sqrt{5}-1}{2}=0.6180339887$；

2）编出计算 $|f(x)|$ 的程序；

3）算出

$$p=a+w(b-a), \quad q=b-w(b-a);$$

4）依 2）算出 $|f(p)|, |f(q)|$；

5）如果在合乎要求的精密度下 $|f(p)|=0$，则停机并打出 p 的数值，同样处理 $|f(q)|$；

6）如果不合于 5）的情形，则比较 $|f(p)|$ 与 $|f(q)|$ 的大小，如果 $|f(p)| \geqslant |f(q)|$，则取 $a_1=a, b_1=p$；如果 $|f(p)|<|f(q)|$，则取 $a_1=q, b_1=b$；

7）如果 b_1-a_1 合乎要求，则停机打出 a_1, b_1 及 $|f(a_1)|$，$|f(b_1)|$；

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

8）不然，回到原来的程序，但以 a_1, b_1 代替 a, b.

注意：如果在 5）停机，我们找到了近似解. 如果在 7）停机，而 $|f(a_1)|$ 不是合乎要求地小，我们可以得出结论：不是问题在 (a_1, b_1) 中无解，就是机器出了毛病. 从 $|f(a_n)|$ 与 $|f(a_{n-1})|$ 的大小，可以来纠正错误.

如果 $f(x) = 0$ 不止一个解，但能估计出每个解所在的范围，用上法可找出所有的解.

§3. 对 开 法

对开法的程序也不复杂，我们可以用来解联立方程

$$f(x, y) = 0, \quad g(x, y) = 0.$$

假定在

$$a \leqslant x \leqslant b, \quad c \leqslant y \leqslant d$$

的范围内有一个解.

1）用黄金分割法算出

$$F\left(x, \frac{1}{2}(c+d)\right) = \left|f\left(x, \frac{1}{2}(c+d)\right)\right|^2 + \left|g\left(x, \frac{1}{2}(c+d)\right)\right|^2, \quad (a \leqslant x \leqslant b)$$

的最小值 p，在 $x = x_1$ 处取这值. 如果在合乎要求的精密度下，$p = 0$，则停机打出 $x_1, \frac{1}{2}(c+d)$；

2）同法算出在 $\left(\frac{1}{2}(a+b), y_1\right)$ 处 $F\left(\frac{1}{2}(a+b), y\right)$ 取最小值 q，如果 q 也合乎精密度且 $q = 0$，则停机打出 $\frac{1}{2}(a+b), y_1$；

3）分四种情形

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

当 $p \geqslant q$, $y_1 > \frac{1}{2}(c+d)$ 时,取

$$a_1 = a,\ b_1 = b,\ c_1 = \frac{1}{2}(c+d),\ d_1 = d;$$

当 $p \geqslant q$, $y_1 < \frac{1}{2}(c+d)$ 时,取

$$a_1 = a,\ b_1 = b,\ c_1 = c,\ d_1 = \frac{1}{2}(c+d);$$

当 $p < q$, $x_1 > \frac{1}{2}(a+b)$时,取

$$a_1 = \frac{1}{2}(a+b),\ b_1 = b,\ c_1 = c,\ d_1 = d;$$

当 $p < q$, $x_1 < \frac{1}{2}(a+b)$ 时,取

$$a_1 = a,\ b_1 = \frac{1}{2}(a+b),\ c_1 = c,\ d_1 = d;$$

4) 以 a_1, b_1, c_1, d_1 代替 a, b, c, d, 重新开始.

附记 $y_1 = \frac{1}{2}(c+d)$ 时, p 不能大于 q, 因为在 $y = \frac{1}{2}(c+d)$ 上不可能取比 p 更小的数值 q, 因而本法仅在 $p = q$, $x_1 = \frac{1}{2}(a+b)$, $y_1 = \frac{1}{2}(c+d)$ 时失效, 也就是无论横向或纵向, $F(x, y)$ 都在 $\left(\frac{1}{2}(a+b), \frac{1}{2}(c+d)\right)$ 处取最小值,为了避免这一情况,我们在 2) 后 3) 前加上一条:

如果 $p = q$, $x_1 = \frac{1}{2}(a+b)$, $y_1 = \frac{1}{2}(c+d)$ 停机打出 a, b, c, d, p 等数值.

这点可能就是解,因为 $\frac{\partial F}{\partial x} = \frac{\partial F}{\partial y} = 0$ 与 $f(x, y) = 0$,

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

$g(x, y)=0$ 等价.

§4. 旋 升 法

与§3同样的问题,旋升法的计算程序如下:

1) 用一维方法算出 $F(x, y)$ 在

$$a \leqslant x \leqslant b, \quad y=y_1=c$$

上的最小值，并设在 $x=x_1$ 时取此值. 如果此值合乎要求，停机,打出 x_1, y_1;

2) 仍用一维的方法算出 $F(x, y)$ 在

$$x=x_1, \quad c \leqslant y \leqslant d$$

上的最小值. 设在 $y=y_2$ 时取此值. 如果此值合乎要求，停机,打出 x_1, y_2;

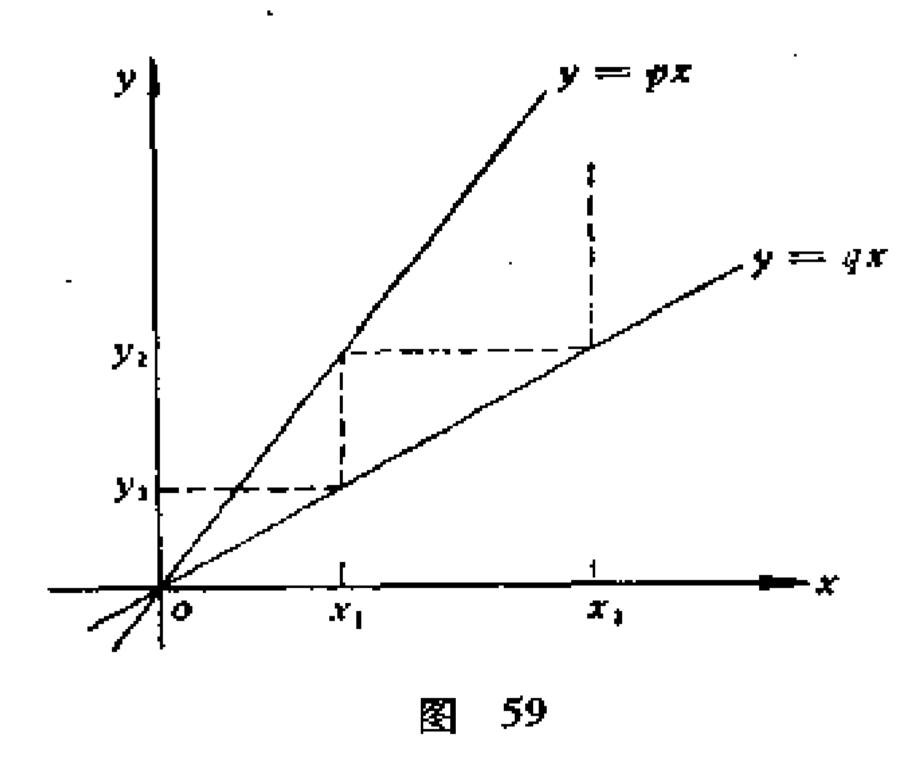


图 59

3) 以 y_2 代替 y_1, 重复以上的计算.

这一方法实际上和以下的方法同一原理: 先求 $f(x, y_1)=0$ 的解 $x=x_1$, 再求 $g(x_1, y)=0$ 的解 $y=y_2$ 等等. 请读者注意，把这方法用到二直线

$$y=px, \quad y=qx$$

$(|p|>|q|)$上去的情形. 若先求 $y_1=qx$ 的解 x_1,后求 $y=px_1$ 的解 y_1,……,便将偏离二直线交点愈来愈远. 但若先求 $y_1=px$ 的解 x_1, 后求 $y=qx_1$ 的解 y_1,……,则能收敛于二直线的交点. 从这一个简单的例子,可以启发出一般原理来,甚至可以启发出收敛的条件来.

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

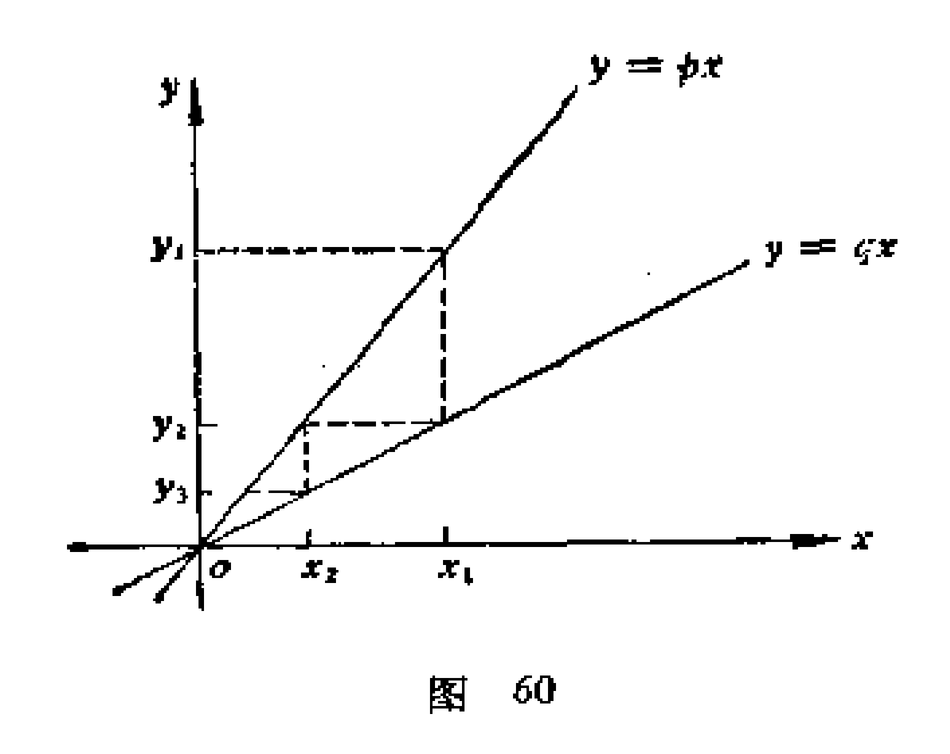


图 60

§5. 数值微分法

我们在研究单变数时，往往用一点及其邻近点的差分来逼近微分，也即

$$f'(x_0) \sim \frac{f(x_0) - f(x_1)}{x_0 - x_1}.$$

在多变数时，必须增加另外的假定，而不能泛泛地用一点及其邻近几点的数值来逼近偏微商．例如在 $x = x_0$，$y = y_0$ 处 $f(x, y)$ 的偏微商 $\frac{\partial f}{\partial x}$，$\frac{\partial f}{\partial y}$ 的值，能否由此点及其两邻点 (x_1, y_1)，(x_2, y_2) 的函数值来逼近呢？不增加条件是不行的．譬如说，三点在一条直线上不行，因为这样只能决定一个方向；夹角太小也不行．另一方面，如果两方向趋于 (x_0, y_0) 的速度太不一致也不行．例如

$$f(x, y) = x^2 + y^2$$

在 $x = 0$，$y = 0$ 时偏微商都等于 0，但 $x = \varepsilon$，$y = \varepsilon^{\frac{2}{3}}$ 时

$$\lim_{\varepsilon \to 0} \frac{x^2 + y^2}{x} = \lim_{\varepsilon \to 0} \frac{\varepsilon^2 + \varepsilon^{\frac{4}{3}}}{\varepsilon} = \infty$$

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

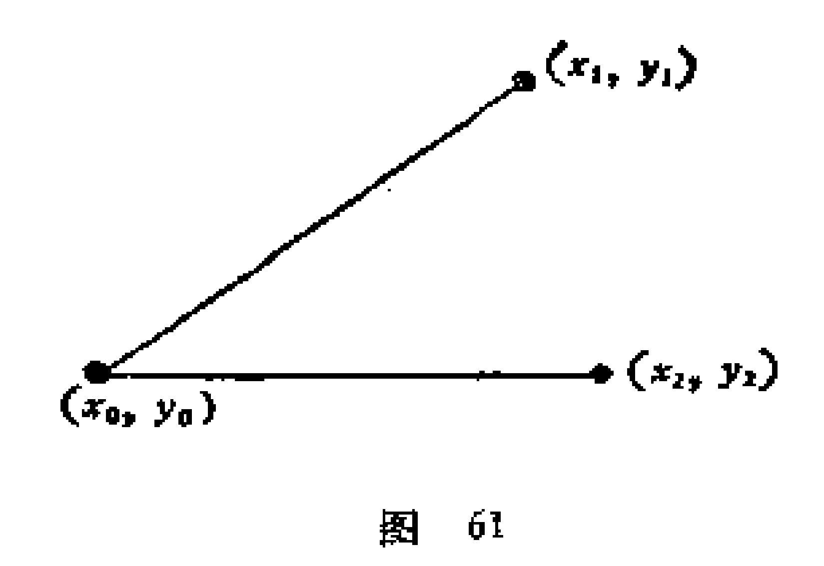


图 61

这是第四(第五)章中我们不是随意取 $s+1$ 点(或 $\frac{1}{2}(s+1)(s+2)$ 点)的道理.

对怎样的三点有此可能呢？实质上，如果面积的绝对值

$$\left\|\begin{matrix} x_1-x_0 & y_1-y_0 \\ x_2-x_0 & y_2-y_0 \end{matrix}\right\| \geqslant K\left(\sqrt{(x_1-x_0)^2+(y_1-y_0)^2}\right.$$

$$\left.+\sqrt{(x_2-x_0)^2+(y_2-y_0)^2}\right)^2.$$

在此条件下，可望成功，读者自证这条件保证了夹角不能任意小.

为了方便起见，还是采用矩阵符号，在 s 维空间给了 $s+1$ 个点 $x^{(i)}(i=0, 1, \cdots, s)$ 从

$$\varphi(x^{(i)}) = a + \boldsymbol{b}x^{(i)} \quad (i=0, 1, 2, \cdots, s) \qquad (1)$$

定出向量 $\boldsymbol{b}$，在什么条件下，向量 $\boldsymbol{b}$ 可以逼近

$$\operatorname{grad}\varphi = \left(\frac{\partial\varphi}{\partial x_1}, \cdots, \frac{\partial\varphi}{\partial x_s}\right)_{x=x^{(0)}}. \qquad (2)$$

定理 1 命 $d_i = \sum_j |x_j^{(i)} - x_j^{(0)}|$ 及

$$M = \begin{pmatrix} x^{(1)} - x^{(0)} \\ x^{(2)} - x^{(0)} \\ \cdots\cdots \\ x^{(s)} - x^{(0)} \end{pmatrix}.$$

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

如果 $M'M$ 的最小特征根 $\geqslant K\sum_{i=1}^{s} d_i^2$，则

$$(b-\mathrm{grad}\varphi(x^{(0)}))(b-\mathrm{grad}\varphi(x^{(0)}))'\leqslant Q\sum_{i=1}^{s} d_i^2.$$

这里 Q 仅依赖于 φ 的二阶微商的上界及 K.

证　由于

$$\varphi(x^{(i)})-\varphi(x^{(0)})=\int_0^1 \frac{d}{d\alpha}\varphi(x^{(0)}+\alpha(x^{(i)}-x^{(0)}))d\alpha$$

$$=\sum_{j=1}^{s}(x_j^{(i)}-x_j^{(0)})\int_0^1 \varphi_{x_j}(x^{(0)}+\alpha(x^{(i)}-x^{(0)}))d\alpha$$

$$=\sum_{j=1}^{s}(x_j^{(i)}-x_j^{(0)})\varphi_{x_j}(x^{(0)})+\sum_{j=1}^{s}(x_j^{(i)}-x_j^{(0)})$$

$$\cdot\int_0^1[\varphi_{x_j}(x^{(0)}+\alpha(x^{(i)}-x^{(0)}))-\varphi_{x_j}(x^{(0)})]d\alpha$$

$$=\mathrm{grad}\,\varphi(x^{(0)})(x^{(i)}-x^{(0)})'$$

$$+\sum_j\sum_k(x_j^{(i)}-x_j^{(0)})(x_k^{(i)}-x_k^{(0)})A_{jk}^{(i)},$$

此处 $A_{jk}^{(i)}$ 仅依赖于 φ 的二阶微商.

由(1)可知

$$\varphi(x^{(i)})-\varphi(x^{(0)})=b(x^{(i)}-x^{(0)})'$$

因此

$$(b-\mathrm{grad}\varphi(x^{(0)}))(x^{(i)}-x^{(0)})'=p_i \tag{3}$$

这儿

$$p_i=\sum_j\sum_k(x_j^{(i)}-x_j^{(0)})(x_k^{(i)}-x_k^{(0)})A_{jk}^{(i)},$$

及

$$|p_i|\leqslant A\left(\sum_j|x_j^{(i)}-x_j^{(0)}|\right)^2=Ad_i^2,$$

K——常数，A 仅依赖于 φ 的二阶微商的上界.

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

(3) 式可以写成为

$$(b-\operatorname{grad}\varphi(x^{(0)}))M'=(p_1,p_2,\cdots,p_s),$$

因此

$$(b-\operatorname{grad}\varphi(x^{(0)}))M'M(b-\operatorname{grad}\varphi(x^{(0)}))'=\sum_{i=1}^{s}p_i^2.$$

由定理的假定可知

$$K\sum_{i=1}^{s}d_i^2(b-\operatorname{grad}\varphi(x^{(0)}))(b-\operatorname{grad}\varphi(x^{(0)})'\leqslant A^2\sum_{i=1}^{s}d_i^4$$

此即所求.

§6. 方程组的数值解

本书第五章所介绍的方法也可以用来处理求方程组的数值解的问题. 求

$$f_1(x_1,\cdots,x_n)=0,\cdots,f_n(x_1,\cdots,x_n)=0 \tag{1}$$

的数值解,用矩阵写法

$$f(x)=0. \tag{2}$$

先给了精密度 η, 例如精确到五位小数, 则取 $\eta=10^{-6}$, 用叠代法

$$x^{(\nu+1)}=x^{(\nu)}-f(x^{(\nu)})A(x^{(\nu)})^{-1}, \tag{3}$$

这儿

$$A(x^{(\nu)})=\begin{pmatrix}\Delta x_1 f(x^{(\nu)})\\ \cdots\cdots\\ \Delta x_n f(x^{(\nu)})\end{pmatrix},$$

$$\Delta x_i f(x)=\frac{f(x+l_i\eta)-f(x)}{\eta}. \tag{4}$$

定理 1 假定

$$\sum_{i=1}^{n}\left|\frac{\partial^2 f_j}{\partial x_i\partial x_k}\right|\leqslant M, \tag{5}$$

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

则

$$\|x^{(\nu+1)}-x^{(\nu)}\| \leqslant \frac{1}{2}M(\|x^{(\nu-1)}-x^{(\nu)}\|^2$$

$$+\eta\|x^{(\nu-1)}-x^{(\nu)}\|)\|A(x^{(\nu)})^{-1}\|. \qquad (6)$$

证 由(3)可知

$$(x^{(\nu+1)}-x^{(\nu)})A(x^{(\nu)})=(f(x^{(\nu-1)})-f(x^{(\nu)})$$

$$+(x^{(\nu)}-x^{(\nu-1)})A(x^{(\nu-1)})$$

$$=\int_0^1\frac{d}{d\alpha}f(x^{(\nu)}+\alpha(x^{(\nu-1)}-x^{(\nu)})d\alpha$$

$$+\sum_{j=1}^{n}(x_j^{(\nu)}-x_j^{(\nu-1)})\Delta x_j f(x^{(\nu-1)})$$

$$=\sum_{j=1}^{n}(x_j^{(\nu-1)}-x_j^{(\nu)})\Big[\int_0^1 f_{x_j}(x^{(\nu)}+\alpha(x^{(\nu-1)}$$

$$-x^{(\nu)}))d\alpha-\int_0^1 f_{x_j}(x^{(\nu-1)}+\alpha l_j\eta)d\alpha\Big]$$

$$=\sum_{j=1}^{n}(x_j^{(\nu-1)}-x_j^{(\nu)})\int_0^1\int_0^1\frac{d}{d\beta}f_{x_j}(x^{(\nu-1)}+\alpha l_j\eta$$

$$+\beta((1-\alpha)(x^{(\nu)}-x^{(\nu-1)})-\alpha l_j\eta)]d\beta d\alpha$$

$$=\sum_{j=1}^{n}(x_j^{(\nu-1)}-x_j^{(\nu)})\int_0^1\int_0^1\sum_{k=1}^{n}[(1-\alpha)$$

$$\cdot(x_k^{(\nu)}-x_k^{(\nu-1)})-\alpha\delta_{jk}\eta]f^{(*)}_{x_jx_k}d\beta d\alpha$$

因此得到

$$\|x^{(\nu+1)}-x^{(\nu)}\|\leqslant\Big\{M\sum_{j=1}^{n}\sum_{k=1}^{n}|x_j^{(\nu-1)}-x_j^{(\nu)}||x_k^{(\nu)}$$

$$-x_k^{(\nu-1)}|\int_0^1\int_0^1(1-\alpha)d\alpha d\beta+\eta M\sum_{j=1}^{n}|x_j^{(\nu)}$$

$$-x_j^{(\nu-1)}|\int_0^1\int_0^1\alpha d\alpha d\beta\Big\}\|A(x^{(\nu)})^{-1}\|\leqslant\frac{M}{2}(\|x^{(\nu)}$$

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

$$- x^{(\nu-1)}\|^2 + \eta\|x^{(\nu)} - x^{(\nu-1)}\|)\|A(x^{(\nu)})^{-1}\|.$$

定理 2 假定

$$\sum_{j=1}^{n}\left|\frac{\partial^2 f_i}{\partial x_j \partial x_k}\right| \leqslant M, \tag{7}$$

$$M\|x^{(\nu)} - x^{(\nu-1)}\|\|A(x^{(\nu-1)})^{-1}\| \leqslant \alpha_{\nu-1} < 1, \tag{8}$$

则得

$$\|A(x^{(\nu)})^{-1}\| \leqslant \frac{1}{1-\alpha_{\nu-1}}\|A(x^{(\nu-1)})^{-1}\|. \tag{9}$$

证 命

$$A(x^{(\nu)}) - A(x^{(\nu-1)}) = (p_{ij}), \quad A(x^{(\nu-1)})^{-1} = (g_{ij}),$$

则

$$\begin{aligned} p_{ij} &= \Delta_{x_j}(f_i(x^{(\nu)}) - f_i(x^{(\nu-1)})) \\ &= \Delta_{x_j}\int_0^1 \frac{d}{d\alpha} f_i(x^{(\nu-1)} + \alpha(x^{(\nu)} - x^{(\nu-1)})d\alpha \\ &= \sum_{k=1}^{n}(x_k^{(\nu)} - x_k^{(\nu-1)})\Delta_{x_j}\int_0^1 f_i x_k(x^{(\nu-1)} \\ &\quad + \alpha(x^{(\nu)} - x^{(\nu-1)}))d\alpha, \end{aligned}$$

因此

$$\begin{aligned} \|I - A(x^{(\nu)})A(x^{(\nu-1)})^{-1}\| &\leqslant \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{l=1}^{n}|p_{ij}||q_{jl}| \\ &\leqslant \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{l=1}^{n}\sum_{k=1}^{n}|x_k^{(\nu)} - x_k^{(\nu-1)}|\int_0^1|\Delta x_j f_i x_k(x^{(\nu-1)} \\ &\quad + \alpha(x^{(\nu)} - x^{(\nu-1)}))|d\alpha|q_{jl}| \\ &\leqslant M\|x_k^{(\nu)} - x_k^{(\nu-1)}\|\|A(x^{(\nu-1)})^{-1}\| \end{aligned} \tag{10}$$

这里用了

$$\sum_i|\Delta x_j f_i x_k| \leqslant \int_0^1\sum_i|f_j x_k x_i|d\beta \leqslant M$$

命

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

$$I - A(x^{(\nu)})A(x^{(\nu-1)})^{-1} = B \tag{11}$$

则

$$A(x^{(\nu)})^{-1} = A(x^{(\nu-1)})^{-1}(I - B)^{-1},$$

由定理的假定及上章 § 1 引理 1 得到所要的结论.

定理 3 如果

$$M\|x^{(1)} - x^{(0)}\|\|A(x^{(0)})^{-1}\| \leqslant \alpha_0 < \frac{3-\sqrt{5}}{2}$$

及

$$\|x^{(\nu)} - x^{(\nu-1)}\| \geqslant \eta,$$

则

$$\|x^{(\nu+1)} - x^{(\nu)}\| < \left(\frac{3+\sqrt{5}}{2}\alpha_0\right)^{2\nu-1}\left(\frac{\sqrt{5}-1}{2}\right)^{\nu}\|x^{(1)} - x^{(0)}\|$$

证明方法和第五章的定理 3 完全一样.

在计算机上使用时，不必验证 (5), (7), (8) 是否适合，而直接应用叠代公式 (3), 如果

$$\|x^{(\nu+1)} - x^{(\nu)}\| < \|x^{(\nu)} - x^{(\nu-1)}\|,$$

即可大胆地做下去，不然，停机换其他方法缩小范围.

我们提请读者注意第五章 § 5 的方法.

§ 7. 求 重 心

在第三章中我们曾提到求重心的问题，高维的重心必须应用多重积分，而多重积分的计算在计算数学中又是一件麻烦的事，一般讲来我们需要计算出 $s+1$ 个 s 重积分

$$V = \int \cdots \int_R dx_1 \cdots dx_s, \quad M_1 = \int \cdots \int_R x_1 dx_1 \cdots dx_s, \quad \cdots,$$

$$M_s = \int \cdots \int_R x_s dx_1 \cdots dx_s.$$

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

这里 R 是一个多面体，也就是适合 p 个不等式

$$R_i: \sum_{j=1}^{s} a_{ij}x_j < b_i \quad (i=1,2,\cdots,p)$$

的点的集合。计算多重积分时，我们用 Mont-Carlo 方法，对这一问题有稍为简化的建议。假定 R 在一个"方体"

$$K: |x_i| < A \quad (i=1,2,\cdots,s)$$

内，利用随机数表(或随机数发生器)在获得 s 个数 $(\xi_1,\cdots,\xi_s)$ 后，看它是否在 R 中。如果不在 R 中就作罢；如果在 R 中则写下

$$1, \xi_1, \xi_2, \cdots, \xi_s; \tag{1}$$

再获得 $\xi_1^{(1)}, \xi_2^{(n)}, \cdots, \xi_s^{(1)}$ 后，如果不在 R 中则作罢，如果在 R 中，则在 (1) 上加上

$$1, \xi_1^{(1)}, \xi_2^{(1)}, \cdots, \xi_s^{(1)};$$

得

$$2, \eta_1^{(1)} = \xi_1 + \xi_2^{(1)}, \cdots, \eta_s^{(1)} = \xi_s + \xi_s^{(1)};$$

再做下去，得 $\xi_1^{(2)}, \cdots, \xi_s^{(2)}$，如果不在 R 中，则作罢，如果在 R 中，则加成

$$3, \eta_1^{(2)} = \eta_1^{(1)} + \xi_1^{(2)}, \cdots, \eta_s^{(2)} = \eta_s^{(1)} + \xi_s^{(2)};$$

$\cdots$；一般是

$$m, \eta_1^{(m-1)} = \eta_1^{(m-2)} + \xi_1^{(m-1)}, \cdots, \eta_s^{(m-1)} = \eta_s^{(m-2)} + \xi_s^{(m-1)}.$$

当 m 充分大时，我们拿

$$\frac{\eta_1^{(m-1)}}{m}, \cdots, \frac{\eta_s^{(m-1)}}{m}$$

来作为重心的坐标。何谓充分大？当然可以有概率建议，我们的简单建议是：如果再做一次和上一次所得的数值差不多就停止(或只要求 $\frac{\eta_1^{(m-1)}}{m}$ 和 $\frac{\eta_1^{(m)}}{m+1}$ 差不多就停止也可以，或要求

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

$$\frac{\eta_1^{(m-1)}+\cdots+\eta_s^{(m-1)}}{m} \text{和} \frac{\eta_1^{(m)}+\cdots+\eta_s^{(m)}}{m}$$

差不多也可以).

这方法的概率误差是极好的，因此可以认为已经符合我们的要求. 也许有人说：概率总是概率，其中包括了万一的失误，但是，这一点小险是要冒的.

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

第三部分　附　录

附录一　度量问题

令 $\varphi(x)$ 是一个不降的函数，而且

$$\varphi(0)=0,\quad \varphi(1)=1.$$

则当 x 过 $[0,1]$ 时，$y=\varphi(x)$ 也从 0 到 1.

在变数 x 上用黄金分割法逐步在

$$x_0,\ x_1,\ x_2,\ \cdots$$

做试验时，等价于 y 在

$$y_0,\ y_1,\ y_2,\ \cdots$$

处做试验，而 $y_n=\varphi(x_n)$.

由中值公式可知

$$y_n-y_{n-1}=\varphi'(\xi)(x_n-x_{n-1}),$$

ξ 在 x_{n-1}, x_n 之间. 如果 $\varphi'(\xi)<1$，则用 y 做变数比用 x 好，不然则不好. 由于 $\varphi'(x)$ 的平均值

$$\int_0^1\varphi'(\xi)d\xi=\varphi(1)-\varphi(0)=1.$$

及 $\varphi'(x)\geqslant 0$，所以 $\varphi'(x)$ 有时大于 1，也有时小于 1.

这说明了一个问题，如果函数 $f(x)$ 的最大值在 x_0，则取 $\varphi'(x_0)<1$ 的 $y=\varphi(x)$ 作为测度为好，但由于我们不知道最大值的所在，因而无法取合适的 $\varphi(x)$. 也就是任何一种测度 $\varphi(x)$ 用黄金分割法，有些 $f(x)$ 上算，有些 $f(x)$ 不上算，也就是用来回调试法，总会碰得巧上算，碰得不好不上算，但上算的可能性极小，因此在一般情况下，我们就用原来普通的测

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

度就行了.

多变数的情形测度的变化影响更大,而且有相互关系,但在实际问题中,一般地用常用的标准度量单位就行了,不要乱改度量单位,但不排斥有理论根据,而选取合适测度. 例如在知道每种变化与几何级数有关时不妨用对数刻度.

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

附录二　与后道工艺过程无关的优选法

在实际中曾经遇到过以下的多因素优选法：要求在区间(a, b)中选取k个点

$$a < x_1 < \cdots < x_k < b$$

使目标函数$f_k(x_1, \cdots, x_k; a, b)$取最大值，但客观上这问题有以下的特性：在$a < x_i$中，$x_1, \cdots, x_{i-1}$也是使某一函数$f_i(x_1, \cdots, x_{i-1}, a, x_i)$取最大值的点，$f_i(x_1, \cdots, x_{i-1}; a, x_i)$与$x_i$以后的点的选择无关.

因为x_1是使$f_2(x_1; a, x_2)$取最大值的点，也就是x_1适合于

$$g_1(x_1, x_2) = \frac{\partial}{\partial x_1} f_2(x_1, a, x_2) = 0;$$

而x_1, x_2也是使$f_3(x_1, x_2; a, x_3)$取最大值的点，也就是适合于

$$g_2(x_1, x_2, x_3) = \frac{\partial}{\partial x_2} f_3(x_1, x_2; a, x_3) = 0;$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

因此变为求如下形式的联立方程的解的问题：

$$\begin{cases} g_1(x_1, x_2) = 0, \\ g_2(x_1, x_2, x_3) = 0, \\ \quad \cdots\cdots\cdots\cdots \\ g_{k-1}(x_1, \cdots, x_k) = 0, \\ g_k(x_1, \cdots, x_k) = 0. \end{cases} \tag{1}$$

先取一个试点$x_1 = x_1^{(1)}$，由(1)的第一式解得$x_2 = x_2^{(1)}$，由第二式解得$x_3 = x_3^{(1)}$，$\cdots$，由第$k-1$个式解得$x_k = x_k^{(1)}$，以

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

$x_1^{(1)}, \cdots, x_k^{(1)}$ 代入 (1) 的最后一式得

$$h_1 = g_k(x_1^{(1)}, \cdots, x_k^{(1)})$$

如果 $h_1 = 0$，则 $x_1^{(1)}, \cdots, x_k^{(1)}$ 就可能是解了，不然不妨假定 $h_1 > 0$，再取 $x_1 = x_1^{(2)}$，同法得

$$h_2 = g_k(x_1^{(2)}, \cdots, x_k^{(2)})$$

如果 $h_2 = 0$，则得解答，如果 $h_2 < 0$，则取 $x_1 = \dfrac{x_1^{(1)} + x_1^{(2)}}{2}$ 进行计算，这样用对分法很快就能得出 (1) 的解的近似值。如果 $h_2 > 0$，则比较 h_1，h_2 哪个大，例如 $h_1 < h_2$，则在 $x_1^{(1)} \to x_1^{(2)}$ 的相反方向选 $x_1^{(3)}$ 就可以找到问题的近似解了。

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

附录三　一致分布点寻优法

从本书一开始就假定了所考虑的函数是单峰的. 在一般工业生产过程中及在理论上较明确的科学实验中，已经在生产中的工艺中，这一假定是可以保证的，这是我们所介绍的优选法能很快地取得大量成果的道理. 但不能排斥，在有些认识不足的探索性的科研问题上会出现非单峰的现象，我们就遇到过"液晶"的配方问题. 如果照"穷举法"或"选优法"做，工作量太大，几辈子都做不完，而且问题不是单峰的，但与一些化学工作者在一起灵活地混合运用了本书所介绍的方法，不但很快地找到了最优点，而且找到了一条"山脊".

在实践经验少，不会抓主要矛盾的条件下，在理论认识不够，不曾获得精确范围的情况下，尝试新的科学研究课题，以下的方法，可起辅助作用.

如果求 k 个变数 $(x_1, \cdots, x_k) = x$ 的函数 $f(x)$ 的最大值，其区域是

$$D:\quad \xi_i \leqslant x_i \leqslant \eta_i \quad (i = 1, 2, \cdots, k).$$

在区域 D 中取一个一致分布的点列

$$x^{(j)} \quad (j = 1, 2, 3, \cdots). \tag{1}$$

在做了若干次试验得出

$$f(x^{(j)}) \quad (j = 1, 2, \cdots, N)$$

后，可得相当满意的函数值. 再在其附近用本书所介绍的优选法.

这一方法的优点在于容易发现复杂的"地形"(非单峰的情况)，但缺点在于虽然比选优法好得多，但比起优选法来次

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

数还嫌太多。

至于一致分布点列的构造，就不在此叙述了。请看华罗庚，王元《数论在近似分析中的应用》(科学出版社，1978). 一致分布点法，在求数值积分上已有好的成效，而求极值的问题可以看成为某一带参数的积分的极限，即

$$\max|f| = \lim_{p\to\infty}\left(\int_0^1\cdots\int_0^1|f(x)|^p dx\right)^{\frac{1}{p}}.$$

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

附录四　求定正方阵的逆

在第二部分第五章 §2 中要用到求定正对称方阵之逆，同时定正方阵求逆法也是种类繁多的多因素优选法的根本依据，因此写这个附录，先总的指明一下，错综复杂的许多方法，实质上都是从解二次方程的"凑平方法"得来的。

1. 求对称定正方阵的逆

$$S=\begin{pmatrix} s_{11}, & \cdots, & s_{1K} \\ \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots \\ s_{k1}, & \cdots, & s_{kk} \end{pmatrix}$$

是一对称方阵，$S=S'$，其第 i 个主对角线方阵是

$$S_i=\begin{pmatrix} s_{11}, & \cdots, & s_{1i} \\ \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots \\ s_{i1}, & \cdots, & s_{ii} \end{pmatrix}$$

又令 $V_{i-1}=(s_{i1},\cdots,s_{i,i-1})$. S_i 的逆 $H_i(=S_i^{-1})$ 可由以下的递归公式得之:

$$S_1^{-1}=H^{(1)}=\frac{1}{s_{11}},$$

$$S_i^{-1}=H^{(i)}=\begin{pmatrix} H^{(i-1)} & 0 \\ 0 & 0 \end{pmatrix}$$

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

$$+\frac{(-V_{i-1}H^{(i-1)},1)'(-V_{i-1}H^{(i-1)},1)}{s_{ii}-V_{i-1}H^{(i-1)}V'_{i-1}}. \qquad (1)$$

因此经 k 次迭代，得 $S^{-1}=H^{(k)}$.

证　由于

$$S_i=\begin{pmatrix}S_{i-1} & V'_{i-1}\\ V_{i-1} & s_{ii}\end{pmatrix}$$

及

$$\begin{pmatrix}I & 0\\ -V_{i-1}S_{i-1}^{-1} & 1\end{pmatrix}\begin{pmatrix}S_{i-1} & V'_{i-1}\\ V_{i-1} & s_{ii}\end{pmatrix}\begin{pmatrix}I & 0\\ -V_{i-1}S_{i-1}^{-1} & 1\end{pmatrix}'$$
$$=\begin{pmatrix}S_{i-1} & 0\\ 0 & s_{ii}-V_{i-1}S_{i-1}^{-1}V'_{i-1}\end{pmatrix}$$

(即"凑平方法")，求逆得

$$S_i^{-1}=H^{(i)}=\begin{pmatrix}I & -H^{(i-1)}V'_{i-1}\\ 0 & 1\end{pmatrix}$$
$$\cdot\begin{pmatrix}H^{(i-1)} & 0\\ 0 & (s_{ii}-V_{i-1}S_{i-1}^{-1}V'_{i-1})^{-1}\end{pmatrix}$$
$$\cdot\begin{pmatrix}I & 0\\ -V_{i-1}H^{(i-1)} & 1\end{pmatrix}=\begin{pmatrix}H^{(i-1)} & 0\\ 0 & 0\end{pmatrix}$$
$$+\frac{1}{s_{ii}-V_{i-1}H^{(i-1)}V'_{i-1}}\begin{pmatrix}H^{(i-1)}V'_{i-1}V_{i-1}H^{(i-1)} & -H^{(i-1)}V'_{i-1}\\ -V_{i-1}H^{(i-1)} & 1\end{pmatrix},$$

即得所证.

附记　求任意方阵 A 的逆也可用此法：由恒等式

$$A^{-1}=A'(AA')^{-1},$$

化为对称方阵 $S=AA'$ 的求逆法. 有些书对 $A=A^{(n,m)}$ 时，将 $A'(AA')^{-1}$ 定义为 A 的广义逆.

2. 线性方程组求解

在第二部分第五章经常遇到线性方程组

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

$$xA = b$$

的求解问题。这里 A 是 n 阶对称正定方阵，b 是 n 维向量，x 则是 n 维解向量。

对于这种方程组的求解，建议采取以下的方法。

第一步：先将系数矩阵 A 表成

$$A = LDL'$$

的形状，这里 L 为一个对角线元素全等于 1 的下三角阵，L' 为其转置，D 为一个对角矩阵。

用 a_{ij}, l_{ij}($1 \leqslant i, j \leqslant n$) 表示矩阵 A, L 中位于第 i 行，第 j 列交点上的元素，自然有

$$a_{ij} = a_{ji},$$

$$l_{ii} = 1, \quad l_{ij} = 0 \quad (1 \leqslant i < j),$$

又用 d_i 表示矩阵 D 的第 i 个对角线元素，于是根据公式

$$l_{ij} = \left(a_{ij} - \sum_{k=1}^{j-1} (l_{ik}d_k) \cdot l_{jk}\right) \Big/ d_j \quad (1 \leqslant j < i \leqslant n)$$

$$d_i = a_{ii} - \sum_{k=1}^{i-1} (l_{ik}d_k) l_{ik} \quad (1 \leqslant i \leqslant n)$$

依次确定出 L 与 D 的元素：

d_1,

… …

$l_{i,1}$, $l_{i,2}$, …, $l_{i,i-1}$, d_i,

… … … …

l_{n1}, l_{n2}, … … … $l_{n,n-1}$, d_n.

在实际计算中，对于一个固定的 i ($1 \leqslant i \leqslant n$)，由于

$$l_{ik}d_k \quad (1 \leqslant k \leqslant i - 1)$$

的反复运用，可以在每算出一个 l_{ik} 后乘以 d_k 而贮存起来，以便后面取用。

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

易见这一部分的计算量(乘除法)的数量级是 $n^3/12$。

第二步，在确定了 L 与 D 后，通过解方程组

$$zL' = b$$

$$yD = z$$

$$xL = y$$

就得出原方程组的解。因为 L, D, L' 都是三角形阵或对角线阵，所以这些方程组都很易解，其乘除法计算量的数量级都不超过 $o(n^2)$。

3. 求抛物体的极值

令

$$z = \alpha + \boldsymbol{p}\boldsymbol{x}' + \frac{1}{2}\boldsymbol{x}S\boldsymbol{x}', \tag{2}$$

此处 $\boldsymbol{p}$ 是 k 维向量，S 是 k 行列的对称定正方阵，"凑平方法"就是以下的方法：

$$z = \alpha - \frac{1}{2}\boldsymbol{p}S^{-1}\boldsymbol{p}' + \frac{1}{2}(\boldsymbol{x}+\boldsymbol{p}S^{-1})S(\boldsymbol{x}+\boldsymbol{p}S^{-1})'$$

$$\geqslant \alpha - \frac{1}{2}\boldsymbol{p}S^{-1}\boldsymbol{p}'. \tag{3}$$

这就是当且仅当

$$\boldsymbol{x} = -\boldsymbol{p}S^{-1} \tag{4}$$

时，z 取这最小值 $\alpha - \frac{1}{2}\boldsymbol{p}S^{-1}\boldsymbol{p}'$

如用 **1.** (1)，则得解方程 (4) 的方法如下，

命 $\boldsymbol{p} = (p_1, \cdots, p_k)$，$\boldsymbol{p}^{(i)} = (p_1, \cdots, p_i)$ 定义

$$\boldsymbol{x}^{(1)} = \boldsymbol{p}^{(1)}H^{(1)}, \cdots$$

$$\boldsymbol{x}^{(i)} = \boldsymbol{p}^{(i)}H^{(i)} = \boldsymbol{p}^{(i)}\begin{pmatrix} H^{(i-1)} & 0 \\ 0 & 0 \end{pmatrix}$$

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

$$+\frac{\boldsymbol{p}^{(i)}(-V_{i-1}H^{(i-1)},1)'(-V_{i-1}H^{(i-1)},1)}{s_{ii}-V_{i-1}H^{(i-1)}V_{i-1}}$$

$$=(x^{(i-1)},0)+\frac{-x^{(i-1)}V'_{i-1}+p_i}{s_{ii}-V_{i-1}H^{(i-1)}V_{i-1}}(-V_{i-1}H^{(i-1)},1),$$

而 $x^{(k)}$ 就是(4)式的解，这就是求(2)式极值的方法.

但这样计算量嫌大，我们可以用 **2.** 的办法去求方程

$$\boldsymbol{x}S=-\boldsymbol{p} \tag{5}$$

的解.

附记 有些书上讨论矛盾方程组

$$\boldsymbol{x}A=\boldsymbol{p},$$

$\boldsymbol{x}$ 是 k 维向量，$\boldsymbol{p}$ 是 l 维向量，$A=A^{(k,l)}$，$k<l$ 的求解问题. 这种方程组本无解，但硬以

$$\boldsymbol{x}=\boldsymbol{p}A'(AA')^{-1}$$

作为其广义解，应用时请注意，这仅对已知 $\boldsymbol{x}$ 与 $\boldsymbol{p}$ 确有线性关系，并且有观测误差时，才可能有用.

如果 $\boldsymbol{x}$ 与 $\boldsymbol{p}$ 并无线性关系，而滥用此法，不但不能正确地反映客观现象，反而导致思想混乱. 例如原来客观问题是在某一范围的内点有极值，如果由若干点的试验数据，而用求广义逆的方法(或线性迴归)来处理，由于线性函数在任一范围的内点是无极值的，很明显这只会导致错误的结论.

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

附录五　离散与连续

1. 在第二部分第五章中所介绍的抛物体法是离散型的方法，如果我们所考虑的函数是易于求出其一、二阶偏微商的，则有以下的连续型的方法.

以 $\boldsymbol{x}=(x_1,\cdots,x_k)$ 表示 k 个变数的向量，假定函数 $f(x)$ 有一、二阶偏微商，以

$$\nabla f(x)=\left(\frac{\partial f}{\partial x_1},\cdots,\frac{\partial f}{\partial x_k}\right)$$

表示 f 的梯度向量，以

$$H_f(x)=\left(\frac{\partial^2 f(x)}{\partial x_i\partial x_j}\right)_{1\leqslant i,\ j\leqslant k}$$

表示 f 的二阶微商矩阵(又称 Hessian 阵)，则由 $f(x)$ 在一点 a 处的幂级数展式的二次以下项

$$f(a)+\nabla f(a)(x-a)'+\frac{1}{2}(x-a)H_f(a)(x-a)' \quad (1)$$

可以得出迭代法. (1) 式的极值可由"凑平方法"得之，如在附录四中所做的那样，若 H_f 为定负矩阵，则 (1) 式就是

$$f(a)+\frac{1}{2}(x-a+\nabla f H_f^{-1})H_f(x-a+\nabla f H_f^{-1})'$$

$$-\frac{1}{2}\nabla f H_f^{-1}(\nabla f)'\leqslant f(a)-\frac{1}{2}\nabla f H_f^{-1}(\nabla f)', \quad (2)$$

即 (1) 式当

$$x=a-\nabla f H_f^{-1} \quad (3)$$

时取极大值. 令 $x=x^{(1)}$ 是一个初始值，则用递推公式

$$x^{(n+1)}=x^{(n)}-\nabla f(x^{(n)})H_f^{-1}(x^{(n)}) \quad (n=1,2,\cdots)$$

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

比第五章更容易证明这方法也是log log型的方法.

2. Davidon-Fletcher-Powell 引进了如下的半离散半连续的方法，从初始值 $x^{(1)}$ 及 $H_1 = I^{(k)}$ 出发，用以下的迭代公式从 $x^{(n)}$, H_n 得 $x^{(n+1)}$, H_{n+1}.

有了 $x^{(n)}$, H_n 后,求单变数 λ 的函数

$$\varphi(\lambda) = f(x^{(n)} - \lambda\nabla(f(x^{(n)}))H_n) \tag{4}$$

的极大值,得 $\lambda = \lambda_n$. 然后取

$$x^{(n+1)} = x^{(n)} - \lambda_n\nabla(f(x^{(n)}))H_n$$

并令

$$y_n = \nabla f(x^{(n+1)}) - \nabla f(x^{(n)}) \tag{5}$$

及

$$H_{n+1} = H_n + \lambda_n\frac{(\nabla f(x^{(n)})H_n)'(\nabla f(x^{(n)})H_n)}{\nabla f(x^{(n)})H_n y_n'} - \frac{(y_nH_n)'(y_nH_n)}{y_nH_ny_n'} \tag{6}$$

关于这个较复杂的收敛性问题,我们只得到

$$\|x^{(0)} - x^{(n+1)}\| = o(\|x^{(0)} - x^{(n)}\|), \tag{7}$$

这里 $x^{(0)}$ 是 $f(x)$ 取极值的点，(7) 式仅仅比

$$\frac{\|x^{(0)} - x^{(n+1)}\|}{\|x^{(0)} - x^{(n)}\|} \leqslant q < 1$$

稍好一些,而后者是属于log 型的.

1) 这方法中要求梯度，如果用差分来代替，则和本书第二部分第二章所介绍的最陡上升法有同样的缺点.

2) $\varphi(\lambda)$ 的极值怎样求? 如果用单因素优选法，一次就要用log log次试验，仅就这一点求法就不可能在数量级上好过抛物体法.

3) 求梯度是连续性的，而求 y_n 是离散性的，如果 $\varphi(\lambda)$ 求极值是用解析方法，则自然也就假定了 $f(x)$ 可以求二阶偏微商,因此 y_n 可以用 $(x^{(n+1)} - x^{(n)})H$ 来代替(相差不多)，这

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

里H是$f(x)$的二阶微商矩阵（Hessian matrix）。

4）不要因为他们只能估计出收敛速度(7)，而以为(7)就是本方法的评价，实质上，用本书的方法是可以估计出比(7)更精密的收敛条件。但不再把时间花在不比抛物体法好得多的问题上，因而作为读者练习(提示：请考虑$\|x^{(0)}-x^{(n+k)}\|$与$\|x^{(0)}-x^{(n)}\|$的关系)。

3. 在 **1.** 中介绍了一种连续型的方法，这类方法无疑可以改进。例如我们还可用以下公式

$$x_i^{(n+1)}=x_i^{(n)}-b_i^j\frac{\partial f(x^{(n)})}{\partial x_j}-\frac{1}{2}b_p^j b_i^s b_q^k\frac{\partial f(x^{(n)})}{\partial x_j}$$
$$\times\frac{\partial f(x^{(n)})}{\partial x_k}\frac{\partial^3 f(x^{(n)})}{\partial x_s\partial x_p\partial x_k}\quad(n=1,2,\cdots)$$

进行迭代，这里上下有相同指标表示从1到k求和，

$$x^{(n)}=(x_1^{(n)}\cdots,x_k^{(n)}),\quad B=\begin{pmatrix}b_1^1\cdots b_k^1\\ \cdots\cdots\\ b_1^k\cdots b_k^k\end{pmatrix}=H(f(x^{(n)}))^{-1}.$$

这个方法比 **1.** 的方法收敛得快些，但还是log log型的方法，且计算量较大，与这方法相应的离散型及半离散半连续型的方法，由于既没有理论上的困难，也没有实际上的显著优越性，这里就不谈了。

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

附录六　几何优选法

1. 光以最短的时间从一点到达另一点．两点之间以直线为最短，因此光从一点 A 到另一点 B 的行程为连 A，B 的直线．

光由一点 A 经镜面上 O 点反射到另一点 B．图中角 α 称为射入角，角 β 称为射出角．如果 $\alpha \neq \beta$，作 B 对镜面的对称点 B'，$OA+OB > AB'$，AB' 交镜面于 O'，则显然有 $OA+OB > O'A+O'B$．因此由光行时间最短推导出射入角 α 一定等于射出角 β．

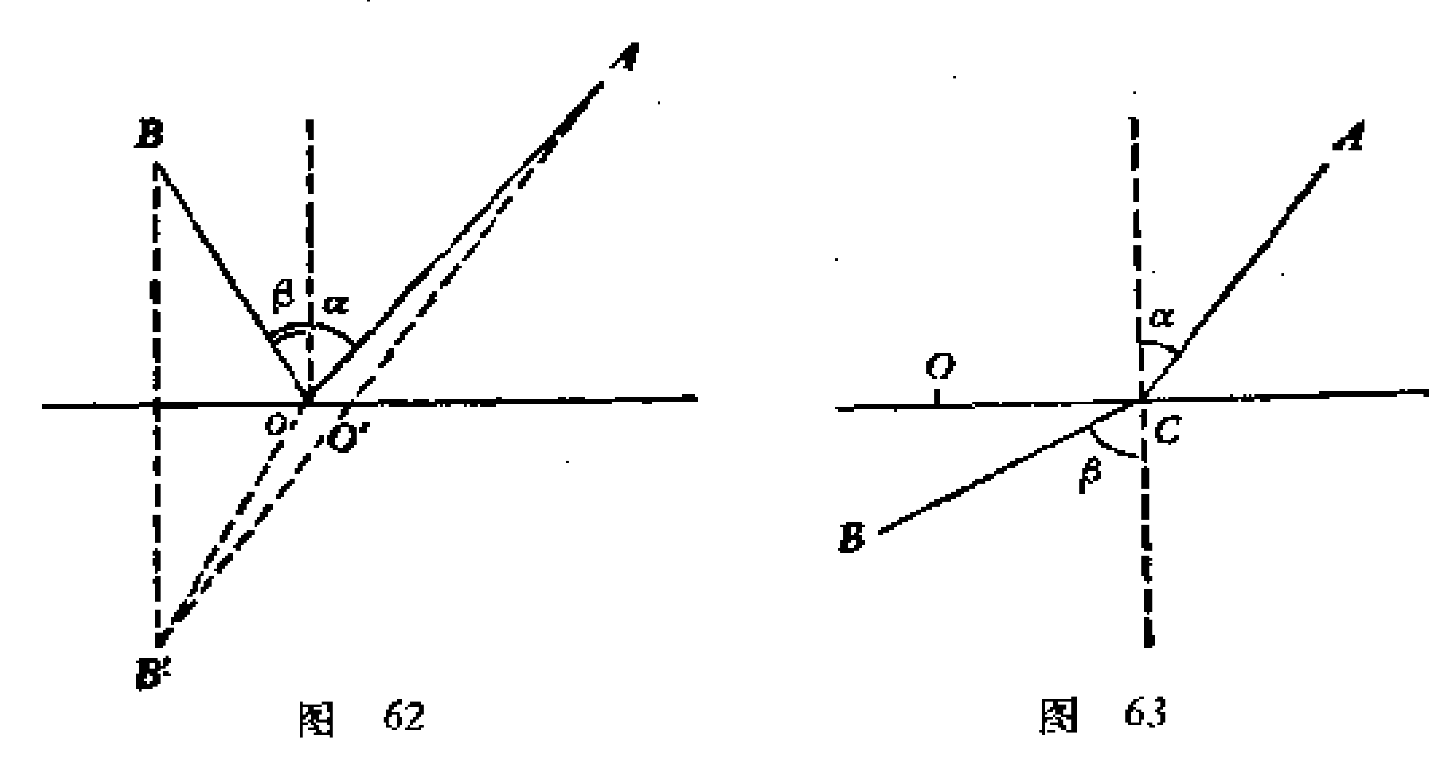


图　62　　　　图　63

光的折射也是这个原理．例如光在空气中以速度 v 进行，在点 C 进入水中以速度 u 达到 B 点，求最速的途径．取水平面上一点 O 作为原点，A，B，C 的坐标各为 (a_1, a_2)，(b_1, b_2)，$(x, 0)$，因此由点 A 到点 B 的时间等于

$$f(x) = \frac{\sqrt{(a_1 - x)^2 + a_2^2}}{v} + \frac{\sqrt{(b_1 - x)^2 + b_2^2}}{u}.$$

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

求 $f(x)$ 的微商使之为 0，

$$\frac{a_1 - x}{v\sqrt{(a_1 - x)^2 + a_2^2}} + \frac{b_1 - x}{u\sqrt{(b_1 - x)^2 + b_2^2}} = 0,$$

因此得

$$\frac{\sin\alpha}{v} = \frac{\sin\beta}{u},$$

即得折射定律.

这个折射定律实际上就是光学设计的主要物理根据. 一个光学镜头(如望远镜,照相机镜头等)通常是由数片透镜组成(图 64). 来自镜头前物面上的光线,按照折射定律依次通过各镜面,在镜头后成像.但这个像不一定是原物体的一个严格的几何相似形,而是有一定的偏差(称为像差). 所谓光学设计，就是在某些约定条件下，通过选择这些透镜的镜面曲率、厚度、间隔、折射率等参数,使所产生的像差尽可能小一些.

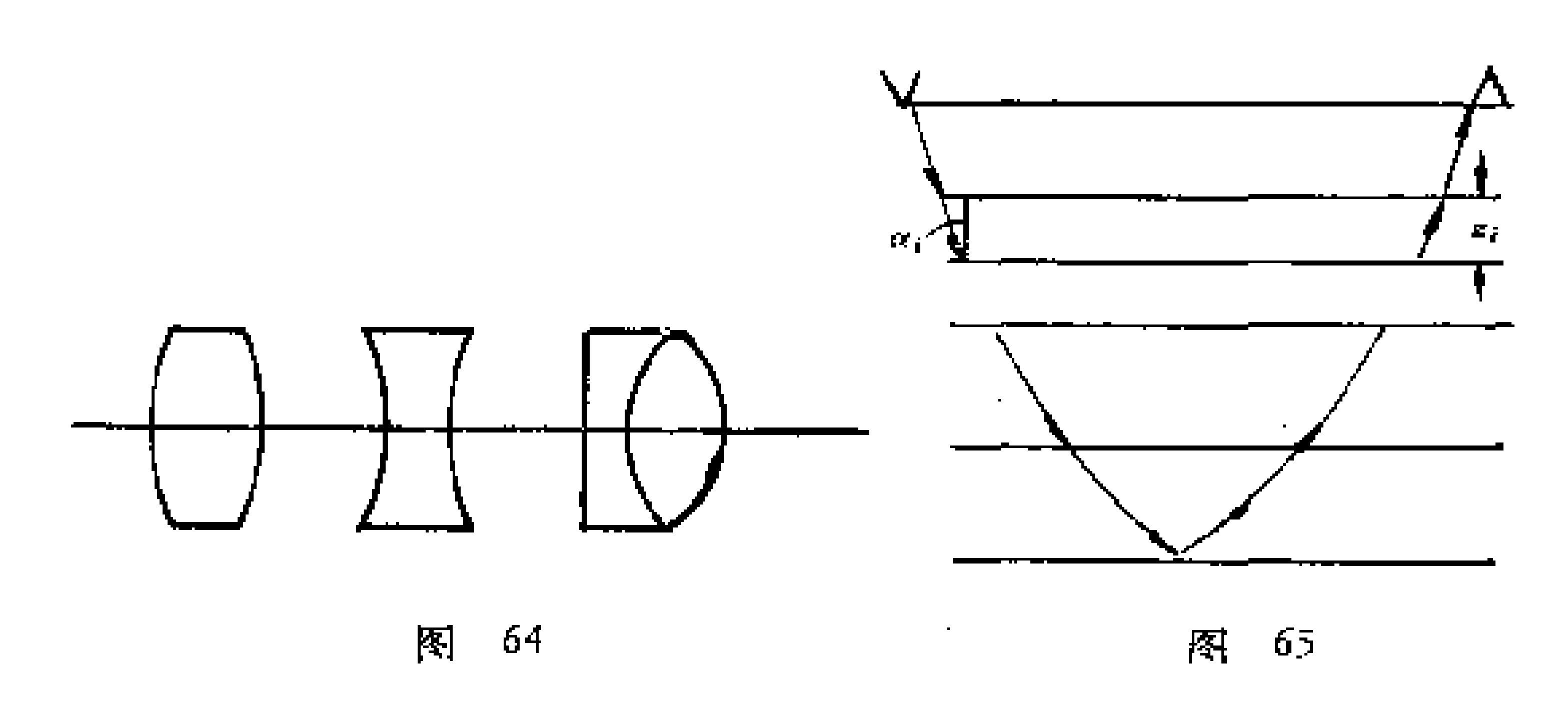


图 64　　　　图 65

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

$$t = 2\sum_{i=1}^{n} \frac{z_i}{v_i \cos\alpha_i} = 2\sum_{i=1}^{n} \frac{z_i}{v_i\sqrt{1-p^2v_i^2}},$$

而距离 x 等于

$$x = 2\sum_{i=1}^{n} z_i \tan\alpha_i = 2\sum_{i=1}^{n} \frac{z_i p v_i}{\sqrt{1-p^2v_i^2}},$$

这里 $p = \dfrac{\sin\alpha_i}{v_i}$, α_i 为第 i 层的波的入射角.

如果地层为连续介质,且速度为深度的函数,则时距之间有如下关系

$$t = 2\int_0^z \frac{dz}{v(z)\sqrt{1-p^2v^2(t)}}, \quad x = 2\int_0^z \frac{pv(z)\,dz}{\sqrt{1-p^2v^2(z)}},$$

这里 z 为反射层的深度.

2. 集散点问题. 在 (x_i, y_i) 点有 a_i 吨原料 $(i=1, 2, \cdots, l)$ 要运到 (ξ, η) 处加工, 加工后又要各运 b_j 吨到 (u_j, v_j) 处 $(j=1, 2, \cdots, m)$. 问选择哪一点 (ξ, η) 运输的吨公里最少? 吨公里等于

$$f(\xi,\eta) = \sum_{i=1}^{l} a_i\sqrt{(x_i-\xi)^2+(y_i-\eta)^2} + \sum_{j=1}^{m} b_j\sqrt{(u_j-\xi)^2+(v_j-\eta)^2}.$$

要解联立方程

$$\frac{\partial f}{\partial \xi} = \sum_{i=1}^{l} \frac{a_i(\xi-x_i)}{\sqrt{(x_i-\xi)^2+(y_i-\eta)^2}} + \sum_{j=1}^{m} \frac{b_j(\xi-u_j)}{\sqrt{(u_j-\xi)^2+(v_j-\eta)^2}} = 0,$$

$$\frac{\partial f}{\partial \eta} = \sum_{i=1}^{l} \frac{a_i(\eta-y_i)}{\sqrt{(x_i-\xi)^2+(y_i-\eta)^2}}$$

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

$$+\sum_{j=1}^{m}\frac{b_j(\eta - v_j)}{\sqrt{(u_j-\xi)^2+(v_j-\eta)^2}}=0$$

中的 ξ, η 的数值并不容易. 1960 年万哲先同志提出了以下的模拟法. (注意: 当 $l=m=1$ 时, 这问题就已比上节光的折射问题难些了.)

先在一块光滑木板上绘制 l 个发点和 m 个收点的分布图, 然后在这些点处各凿一光滑小洞. 在板面上置一光滑金属小圆环, 小圆环上系 $l+m$ 根绳子, 每条绳子各穿过一个小洞, 通过发点 (x_i, y_i) 的绳子系上质量为 a_i 的重物, 通过收点 (u_j, v_j) 的绳子系上质量为 b_j 的重物. 这样当小圆环达到平衡位置时, 平衡位置就是所找的 (ξ, η).

我们引入平面向量

$$\boldsymbol{e}_i=\left(\frac{\xi-x_i}{\sqrt{(x_i-\xi)^2+(y_i-\eta)^2}},\right.$$

$$\left.\frac{\eta-v_i}{\sqrt{(x_i-\xi)^2+(y_i-\eta)^2}}\right),$$

$$\boldsymbol{e}_j'=\left(\frac{\xi-u_j}{\sqrt{(u_j-\xi)^2+(v_j-\eta)^2}},\right.$$

$$\left.\frac{\eta-v_j}{\sqrt{(u_j-\xi)^2+(v_j-\eta)^2}}\right),$$

则上面的联立方程就是

$$\sum_{i=1}^{l}a_i\boldsymbol{e}_i+\sum_{j=1}^{m}b_j\boldsymbol{e}_j'=0,$$

这正是 $l+m$ 个力的平衡方程, 也就是上述模拟法的理论根据.

3. 邮路问题. 一个邮递员每次上班, 要走遍他负责送信的每条路, 然后回到邮局. 问应该怎样走才能使他所走的路程最短? 这就是邮路问题.

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

首先考虑这样一个问题. 如下面这样一个道路图(每条线上的数字表示距离,小圆圈中的数字表示交叉路口的编号). 我们从一点出发,有没有可能沿每条路都走过一遍(不许重复也不许遗漏),然后回到出发点?

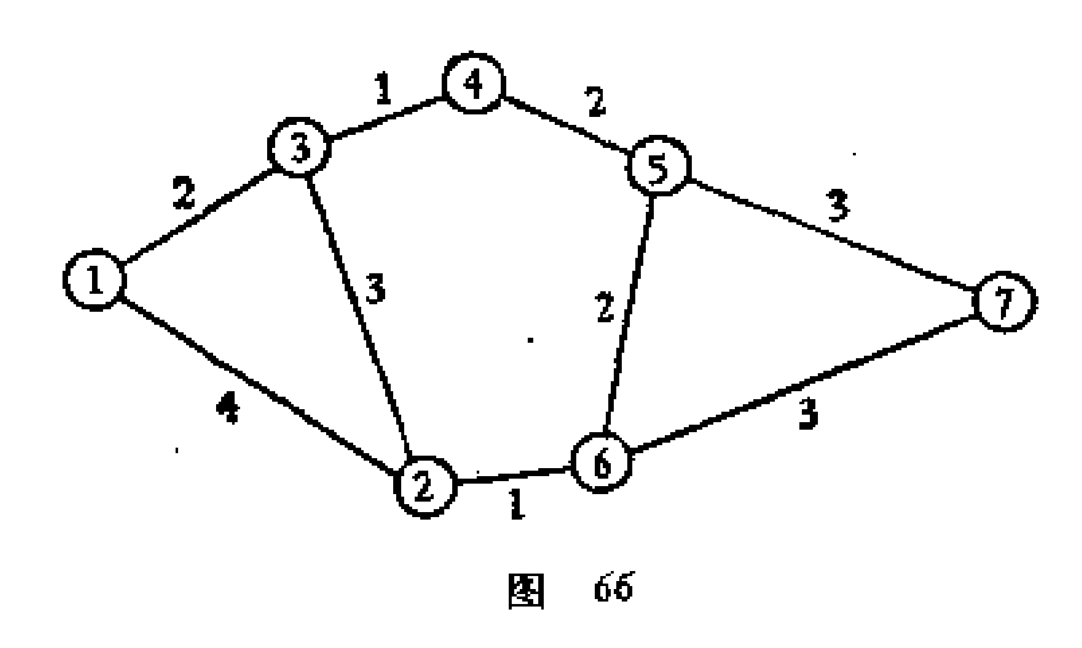


图 66

我们把交叉路口分为两类,偶数条路的交叉点称为偶点,奇数条路的交叉点称为奇点. 例如上图中,①,④,⑦ 是偶点,②,③,⑤,⑥ 是奇点. 易见,当一个图中只要有一个奇点时,就不可能那样走一遍了,因为对一个路口来说,有走进去的一条路,必有走出来的一条路. 反之也不难证明,当道路图上没有奇点时便可以那样走一遍了. 如果起点和终点可以不同,则可以允许有二个奇点. 这就是所谓一笔画问题.

回到邮路问题,如果道路图上没有奇点,问题就已经解决了. 如果有奇点,我们总可以通过在某一些路上再重复走一遍的办法来消灭奇点,这相当于在图上再添上一些路. 上图中,如果在 ② 与 ③ 之间,⑤ 与 ⑥ 之间各添上一条路(即在 ②—③,⑤—⑥ 上走两次),奇点就没有了,于是就可以一笔画了. 但我们还要求所走的路程最短,因此问题归结为:如何选择要重复的路程,既能消灭奇点,又能使重复的路程最短?

这个问题已被管梅谷同志解决. 他证明了:只要在每个

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

圈上重复的路程不超过整个圈长的一半，这时就可以得到路程最短的走法了.

例如在上页图中，如果重复 ②—③，⑤—⑥ 两条路，可以消灭奇点，但在圈 ②—③—④—⑤—⑥—② 上重复路程 $2+3=5$ 已超过整个圈长之半 (4.5)，所以它不是路程最短的走法. 但我们可以改为重复 ②—⑥，③—④，④—⑤，这样既消灭了奇点，又使得每个圈上重复路程都不超过整个圈长之半，于是得到这样一个路程最短的走法:

①→③→④→⑤→⑥→⑦→⑤→④→③→②→⑥→②→①

综合上述，得到这样一个解决邮路问题的步骤: 首先把道路图上所有的奇点找出来，然后选择一些需要重复的路，使奇点消灭. 再在每个圈上检查重复路程是否超过整个圈长之半，如有超过的情况，则在这个圈上，把原来不重复的路定为重复的路，而原来重复的路定为不重复的路. 这样逐步调整，当所有圈上重复路程都不超过整个圈长之半时，就达到路程最短了. 最后将整个图形一笔画出.

一般来说一个图上所具有的圈数是很多的，是否必要每个圈都检查? 如上图共有六个圈，实际上只要检查三个圈就够了. 一般，如一个图有 s 条路，n 个点，则需要检查的圈数为 $s-n+1$.

在画交通图时还要注意一个问题，一条很宽的马路，实际上是需要在马路的两侧各走一次，这时在交通图上要画两条路，而不是画一条路.

这里提出的问题，不只是邮递员会碰到，在其他一些场合也会碰到，如马路上扫地、喷水的车子所走的路线.

4. 蜂房问题. 我们研究与蜂房结构有关的一个数学问题. 从正面看，蜂房是一个正六边形，但它并不是一个六棱

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

柱，它的底部是由三个相同的菱形拼成的（图 67）．菱形的钝角为 109°28′，锐角为 70°32′．说得更具体些，拿一支六棱柱的铅笔，未削之前，铅笔一端的形状是 $ABCDEF$ 正六角形（图 68），通过 AC 一刀切下一角，把三角形 ABC 搬置 AOC 处，过 AE，CE 切如此同样两刀，所堆成的形状就是一个蜂房．一个蜂巢就是两排这样的蜂房，底部和底部相接而成．

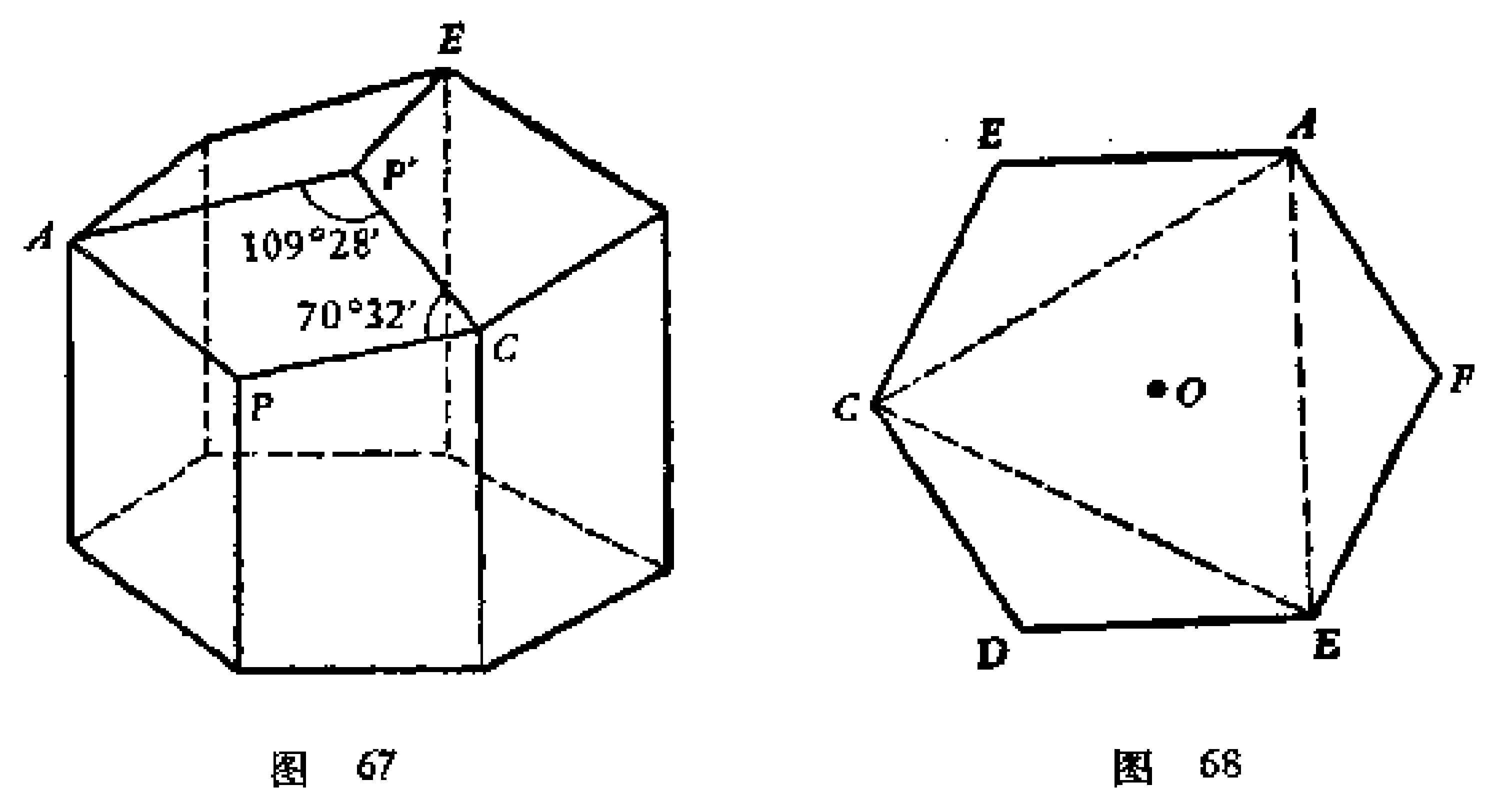


图 67　　　　图 68

建造一个蜂房所消耗的材料是与表面积成正比的．我们问：怎样切出来使所拼成三个菱形做底的六棱柱的表面积最小．

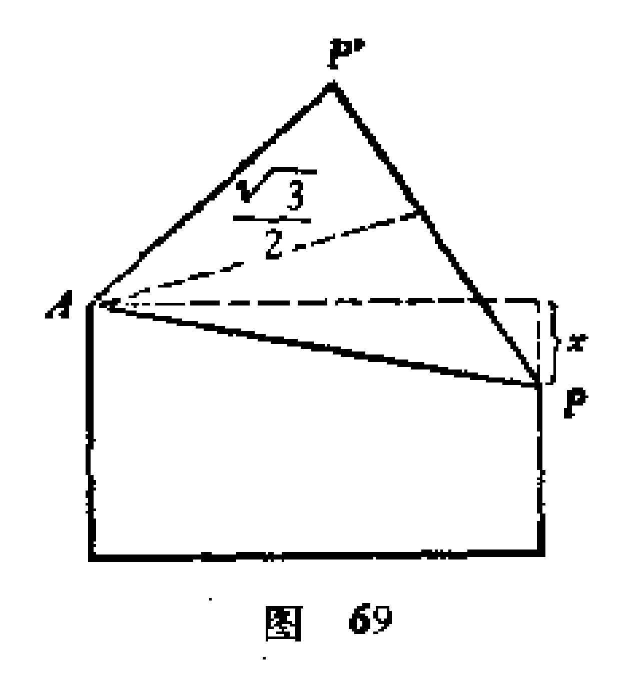


图 69

假定六棱柱边长是 1，这时 $AC=\sqrt{3}$．把图 67 的尖顶六棱柱表面分成六份，把其中之一摊平下来，得出图 69 的形状．从宽为 1 的长方形切去一角，切割处成边 AP，以 AP 为腰，$\frac{\sqrt{3}}{2}$ 为高作等腰三角形，假定被切去的三角形高为 x，则

$$AP=\sqrt{1+x^2},$$

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

$$PP' = 2\sqrt{1 + x^2 - \left(\frac{\sqrt{3}}{2}\right)^2} = \sqrt{1 + 4x^2},$$

$\triangle APP'$ 的面积为

$$\frac{\sqrt{3}}{4}\sqrt{1 + 4x^2},$$

所以问题变为求 x，使

$$f(x) = -\frac{1}{2}x + \frac{\sqrt{3}}{4}\sqrt{1 + 4x^2}$$

最小.

我们用一个初等方法来求 $f(x)$ 的极小值. 令 $2x = t - \frac{1}{4t}$ $(t > 0)$，则因算术平均 $\geqslant$ 几何平均，可见

$$\begin{aligned} f(x) &= -\frac{1}{4}\left(t - \frac{1}{4t}\right) + \frac{\sqrt{3}}{4}\left(t + \frac{1}{4t}\right) \\ &= \frac{\sqrt{3} - 1}{4}t + \frac{\sqrt{3} + 1}{4} \cdot \frac{1}{4t} \\ &\geqslant 2\sqrt{\frac{(\sqrt{3} - 1)(\sqrt{3} + 1)}{4^3}} = \frac{1}{\sqrt{8}} \end{aligned}$$

而且仅当

$$\frac{\sqrt{3} - 1}{4}t = \frac{\sqrt{3} + 1}{4} \cdot \frac{1}{4t}$$

时取等号，这时 $t = \frac{1 + \sqrt{3}}{2\sqrt{2}}$，而

$$x = \frac{1}{2}\left(\frac{1 + \sqrt{3}}{2\sqrt{2}} - \frac{2\sqrt{2}}{4(1 + \sqrt{3})}\right) = \frac{1}{\sqrt{8}},$$

所以 $x = \frac{1}{\sqrt{8}}$ 时，$f(x)$ 取极小值 $\frac{1}{\sqrt{8}}$.

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

令$\angle PAP' = \gamma$，由余弦公式得到

$$2(1+x^2)\cos\gamma = 2(1+x^2)-(1+4x^2) = 1-2x^2,$$

以$x=\dfrac{1}{\sqrt{8}}$代入得到

$$\cos\gamma = \frac{1-2x^2}{2(1+x^2)} = \frac{3}{8}\Big/\left(1+\frac{1}{8}\right) = \frac{1}{3},$$

因此得到$\gamma = 70°32'$，与蜂房的实际角度相一致。我们进一步问：蜂房是否也具有"在体积一定时表面积最小"的特性？

下面将证明：一个形如蜂房的尖顶六棱柱，体积V为一定时，表面积(不算底面)最小值为$3\sqrt{2}\,V^{2/3}$，而且当六角形边长是$\sqrt{\dfrac{2}{3}}\,V^{1/3}$，高度是$\dfrac{\sqrt{3}}{2}\,V^{1/3}$时取最小值。

以边长a的正六角形为底，以b为高的六棱柱，其六个顶点顺次为$ABCDEF$，如上述，过B(或D或F)棱$\dfrac{1}{\sqrt{8}}a$处$\left(假定b>\dfrac{1}{\sqrt{8}}a\right)$及$A$，$C$(或$C$，$E$或$E$，$A$)作一平面，切下三个四面体，反过来堆在顶上，就得到一个以三个菱形做底的尖顶六棱柱。易见它的体积V和表面积S为：

$$V = \frac{3\sqrt{3}}{2}a^2b,$$

$$S = 6a\left(b+\frac{a}{\sqrt{8}}\right).$$

我们有

$$S = \frac{4V}{\sqrt{3}a}+\frac{3}{\sqrt{2}}a^2 = \frac{2V}{\sqrt{3}a}+\frac{2V}{\sqrt{3}a}+\frac{3}{\sqrt{2}}a^2$$

$$\geqslant 3\left(\frac{2V}{\sqrt{3}a}\cdot\frac{2V}{\sqrt{3}a}\cdot\frac{3}{\sqrt{2}}a^2\right)^{1/3} = 3\sqrt{2}\,V^{2/3},$$

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

而且仅当

$$\frac{2V}{\sqrt{3}a}=\frac{3}{\sqrt{2}}a^2,\quad a=\sqrt{\frac{2}{3}}V^{1/3}$$

时 S 取最小值，这时

$$b=\frac{2V}{3\sqrt{3}a^2}=\frac{1}{\sqrt{3}}V^{1/3}>\frac{a}{\sqrt{8}}.$$

尖顶六棱柱高度为

$$b+\frac{1}{\sqrt{8}}a=\frac{1}{\sqrt{3}}V^{1/3}+\frac{1}{2\sqrt{3}}V^{1/3}.$$

即上述结论得证.

通过实测，蜂房的大小与上例并不一致 $\Big($经过实测，$a\doteq0.35$cm，高为 0.70cm，而按上面的结论，高应是 $b+\frac{1}{\sqrt{8}}a=\frac{1}{\sqrt{2}}a+\frac{1}{\sqrt{8}}a=\frac{3}{\sqrt{8}}a\doteq0.38cm\Big)$，所以蜂房并不具有“在体积一定时表面积最小”的特性．而蜜蜂是根据自己身材设计，用材最少的窠.

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

附录七　多目标问题

在实际中经常会遇到多目标的问题，也就是我们要优选一些参数，使几个目标函数同时都达到较好的水平。这里我们建议这样一个方法，在这些指标中，根据当前的情况，选择一个优先要考虑的指标作为我们的目标，而其他的指标希望它不要超出某个界限，也就是引入一些不等式的约束条件，这样，问题就化成单目标的问题了.

有些实际问题可以把多目标问题化为单目标问题.例如，用一个关于经济价值的指标把原来关于劳力、各种材料消耗等的指标合为一个.

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

附录八 重复试验

在得出较好的方案后，按照这个方案固定所有因素反复做试验有时是必要的，但可能会得出不完全相同的数据，对这些数据怎样处理？如果做的试验次数不太多，例如只做几个，十几个或几十个，可用以下的方法：

把试验数据按次序排列起来

$$x_1 \leqslant x_2 \leqslant \cdots \leqslant x_n,$$

得结论：就做几次试验的结果来说，100% 落在范围 $[x_1, x_n]$ 之间，落在 $[x_2, x_{n-1}]$ 之间的可能性是 $(n-2)/n$ 等等.

如果做的试验次数实在多，可用以下的方差分析法：先求平均数

$$\bar{x} = \frac{1}{n}(x_1 + \cdots + x_n)$$

及方差

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(\bar{x} - x_i)^2,$$

而结论是：在 $\left(\bar{x} - \frac{S}{\sqrt{n}}, \bar{x} + \frac{S}{\sqrt{n}}\right)$ 之间的可能性约为 68%；在 $\left(\bar{x} - \frac{2S}{\sqrt{n}}, \bar{x} + \frac{2S}{\sqrt{n}}\right)$ 之间的可能性约为 95% 等等.（**附记** 为什么 S^2 的分母是 $n-1$，因其分子为

$$\sum_{i=1}^{n} x_i^2 - n\bar{x}^2 = \sum_{i=1}^{n} x_i^2 - \frac{1}{n}(x_1 + \cdots + x_n)^2.$$

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

这是 $x_1, \cdots, x_n$ 的二次型，它的方阵

$$I - \frac{1}{n}(1, \cdots, 1)'(1, \cdots, 1)$$

的秩显然为 $n-1$，也就是实际上 S^2 的分子不是 n 个数的平方和而是 $n-1$ 个数的平方和)。

如果得出两种较好的方案，各自固定所有因素后反复做试验，可能会各自得出不完全相同的试验数据，怎样比较这两种方案的优劣？如果做的试验次数不太多，则可将所有数据按大小次序排列起来，按本书第一部分第三章 § 9 所介绍的办法来判断。如果做的试验次数实在多，可用以下的方差分析法：若第一种方案做了 n_1 次试验，得数据为 $x_{11}, \cdots, x_{1n_1}$；第二种方案做了 n_2 次试验，得数据为 $x_{21}, \cdots, x_{2n_2}$。分别求平均数

$$\bar{x}_1 = \frac{1}{n_1}\sum_{i=1}^{n_1} x_{1i}, \quad \bar{x}_2 = \frac{1}{n_2}\sum_{i=1}^{n_2} x_{2i}$$

及方差

$$S_1^2 = \frac{1}{n_1-1}\sum_{i=1}^{n_1}(\bar{x}_1 - x_{1i})^2, \quad S_2^2 = \frac{1}{n_2-1}\sum_{i=1}^{n_2}(\bar{x}_2 - x_{2i})^2.$$

若 $|\bar{x}_2 - \bar{x}_1| > 1.645\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}$，则约有 90% 的可能性，这两种方案所得的结果是有差异的；若

$$|\bar{x}_2 - \bar{x}_1| > 1.96\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}},$$

则约有 95% 的可能性，这两种方案所得的结果是有差异的；等等。

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

附录九 0-1 变元法

1. 在优选法形成一数学分支之前，曾有一些方法，科学根据不太足，先验地假定某一参数仅在若干个水平中选取，最简单的是假定在两个水平(或三个水平)中选取，非此即彼。他们所用的数学工具，与其说用了连续变元，不如说是用了 0-1 变元，即取不取变元，以两个水平为例，温度或取 300℃ 或取 600℃。我们不妨说取 300℃ 时这变元的值是 1，不取 300℃ 时，这变元的值是 0。换变数 $\eta=\dfrac{600-\xi}{600-300}$ 就变为 0-1 变元 η 了。下面，我们简略介绍一下，并且从原则上更提高一些。

一个参数 ξ 限定在 n 个水平

$$\xi_1, \xi_2, \cdots, \xi_n$$

中选取其中之一，可以用 n 个 0-1 变元 $x_1, \cdots, x_n$ 来代表，这些 $x_1, \cdots, x_n$ 只取值 0 或 1，而且

$$x_1+\cdots+x_n=1.$$

即 n 个水平中每次只能取一个。

如果某一目标函数 T，依赖于 k 个参数 $\xi^{(1)}, \cdots, \xi^{(k)}$，而 $\xi^{(i)}$ 有 n_i 个水平

$$\xi_1^{(i)}, \cdots, \xi_{n_i}^{(i)},$$

其对应的 0-1 变元 $x_1^{(i)}, \cdots, x_{n_i}^{(i)}$，假定其目标函数 T 是最简单的一次模型(这种假定的可靠性极小)，

$$T=\sum_{i=1}^{k}\sum_{j=1}^{n_i}\alpha_j^{(i)}x_j^{(i)}, \tag{1}$$

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

$$\sum_{j=1}^{n_i} x_j^{(i)} = 1, \quad x_j^{(i)} = 0 \text{ 或 } 1. \tag{2}$$

用

$$x_{n_i}^{(i)} = 1 - \sum_{j=1}^{n_i-1} x_j^{(i)}$$

代入 (1) 则得

$$T = \sum_{i=1}^{k} \left(\sum_{j=1}^{n_i-1} \alpha_j^{(i)} \cdot x_j^{(i)} + \alpha_{n_i}^{(i)} \left(1 - \sum_{j=1}^{n_i-1} x_j^{(i)} \right) \right)$$

$$= \sum_{i=1}^{k} \alpha_{n_i}^{(i)} + \sum_{i=1}^{k} \sum_{j=1}^{n_i-1} (\alpha_j^{(i)} - \alpha_{n_i}^{(i)}) x_j^{(i)}. \tag{3}$$

这式子共有 $n_1 + \cdots + n_k - k$ 个独立的 0-1 变元 $x_j^{(i)}$，但必须适合

$$\sum_{j=1}^{n_i-1} x_j^{(i)} \leqslant 1. \tag{4}$$

在 k 个参数 $\xi^{(i)}$ 中，每一参数取一个水平的情况做一次试验，也就是取一个 0-1 变元组 $x_j^{(i)}$，得出一个 T 来，做了 N 次试验，得出 N 个以

$$\sum_{i=1}^{k} \alpha_{n_i}^{(i)}, \quad \alpha_j^{(i)} - \alpha_{n_i}^{(i)} \quad (1 \leqslant i \leqslant k, 1 \leqslant j \leqslant n_i - 1) \tag{5}$$

为待定系数的一次联立方程组. 待定系数总共有 $n_1 + \cdots + n_k - k + 1$ 个. 因而做

$$N = n_1 + \cdots + n_k - k + 1 \tag{6}$$

次试验，就可以从解一次联立方程中得出 (5) 的数值. 但请注意，我们所做的试验，要使所用到的行列式不等于 0，定出 (5) 后，就用以下的方法判断参数 $\xi^{(i)}$ 该取哪一个水平. 这方法是，如果对一个参数 $\xi^{(i)}$，

$$\alpha_j^{(i)} - \alpha_{n_i}^{(i)} \quad (j = 1, \cdots, n_i - 1) \tag{7}$$

都是负的，则取 $x_{n_i}^{(i)} = 1$，即 $\xi^{(i)}$ 的第 n_i 个水平 $\xi_{n_i}^{(i)}$；如果不

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

然，而 (7) 中以 $j = j_0(i)$ 为最大，则取 $x_{j_0(i)}^{(i)} = 1$，即取水平 $\xi^{(i)} = \xi_{j_0(i)}^{(i)}$。因此按模型，$T$ 在

$$\xi^{(1)} = \xi_{j_0(1)}^{(1)}, \cdots, \xi^{(k)} = \xi_{j_0(k)}^{(k)}$$

处取最大值．因而他们就得出结论，这一点是目标函数 T 的最优点了．当然模型是很不可靠的，因为线性模型是没有最大值的，有些同志认识到这方法的不可靠性，把这一点称为有参考性的最优点，也就是承认了这一方法不属于有收敛性保证的优选法的范围．

2. 为了更明确地说明问题，我们以每个参数仅取两个水平为例．例如参数 $\xi^{(i)}$ 只取两个水平 $\xi_1^{(i)}$ 或 $\xi_2^{(i)}$ 之一．换变数

$$x_i = \frac{\xi_2^{(i)} - \xi^{(i)}}{\xi_2^{(i)} - \xi_1^{(i)}},$$

则当取第一水平 $\xi_1^{(i)}$ 时，$x_i = 1$，当取第二水平时（即当不取第一水平时），$x_i = 0$．如果假定目标函数原来是

$$T = \alpha_0 + \sum_{i=1}^{k} \alpha_i \xi^{(i)},$$

也就完全等同于目标函数为

$$T = \beta_0 + \sum_{i=1}^{k} \beta_i x_i$$

的 0–1 变元 x_i 的问题了．

由于基本假定易于失真，我们就不再烦琐地叙述多种多样选取 $x_i (i = 1, \cdots, k)$ 做试验的方法．更不谈如何用统计方法来分析其试验结果了．要之，本方法不但不能保证得到最优点，甚至于不能保证自己先验预定的 $n_1, \cdots, n_k$ 个点中的最优点．

至于多做几次试验，即上节中取 $N > n_1 + \cdots + n_K - K + 1$，用附录四中所提到的求矩阵的广义逆法来解矛盾方程，那就更不对头了．

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

例 1 取 $(x_1, \cdots, x_k) = \mathbf{0}$, $\boldsymbol{e}_1, \cdots, \boldsymbol{e}_i, \cdots, \boldsymbol{e}_k$, 这里 $\boldsymbol{e}_i$ 是一个 k 维向量，除第 i 分量是 1 外，其它诸支量都等于 0. 所对应的 $k+1$ 方阵是

$$\begin{pmatrix} 1 & 0 & \cdots\cdots & 0 \\ 1 & 1 & \cdots\cdots & 0 \\ 1 & 0 & 1 \quad \cdots & 0 \\ \vdots & 0 & \ddots & \\ \vdots & \vdots & \ddots & \\ 1 & 0 & \cdots\cdots & 1 \end{pmatrix}$$

这一方法对应于从一点出发，向每一方向都试探一步而后定上升的方向，很明显这并不比一步一升高的"瞎子爬山法"更好些.

例 2 考虑某一配方由三种化学药品所配成，每种药品所取的用量有三种水平，混合后加温，温度也是有三个水平，这是一个四个参数的问题，三个参数表示用量，一个参数表示温度，因此要求做 $3\times4-4+1=9$ 次试验，就能定出其线性目标函数，但这一数学结构易于失真，所以也就不再讨论这批试验点的安排问题，更不考虑不分主次，把所有的因素善同地安排实验的方法了.

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

参 考 文 献

[1] D. J. Wilde, Optimum seeking methods (1964).

[2] D. J. Wilde, C. S. Beightler,Foundations of optimization (1967).

[3] G. S. G. Bevridge, R. S. Schechter, Optimization: theory and practice (1970).

[4] J. W. Blakenuse, S. H. Davis, Optimization techniques (1964).

[5] M. J. D. Powell, An efficient method for finding the minimum of a function of several variables without calculating derivatives, *Computer J.*, 7 (1964), 155—162.

[6] R. Fletcher., M. J. D. Powell, A rapidly convergent descent method for minimization, *Computer J.*, **6** (1963), 163—168.

[7] J. Kiefer, Sequential minimax search for a maximum, *Proc. Amer. Math. Soc.*, **4** (1953), 502—506.

[8] W. Sperdley, Optimization (1969).

[9] A. Lavi, T. P. Vogl, Recent advances in optimization techniques (1966).

1970 年后国内各地出版的优选法文献和推广优选法成果资料汇编已达数百种，内容丰富，因总数浩繁，此处不能一一收录。

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

参 考 文 献

[1] D. J. Wilde, Optimum seeking methods (1964).

[2] D. J. Wilde, C. S. Beightler, Foundations of optimization (1967).

[3] G. S. G. Bevridge, R. S. Schechter, Optimization: theory and practice (1970).

[4] J. W. Blakenuse, S. H. Davis, Optimization techniques (1964).

[5] M. J. D. Powell, An efficient method for finding the minimum of a function of several variables without calculating derivatives, *Computer J.*, 7 (1964), 155—162.

[6] R. Fletcher, M. J. D. Powell, A rapidly convergent descent method for minimization, *Computer J.*, **6** (1963), 163—168.

[7] J. Kiefer, Sequential minimax search for a maximum, *Proc. Amer. Math. Soc.*, **4** (1953), 502—506.

[8] W. Sperdley, Optimization (1969).

[9] A. Lavi, T. P. Vogl, Recent advances in optimization techniques (1966).

1970 年后国内各地出版的优选法文献和推广优选法成果资料汇编已达数百种，内容丰富，因总数浩繁，此处不能一一收录。

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

参 考 文 献

[1] D. J. Wilde, Optimum seeking methods (1964).
[2] D. J. Wilde, C. S. Beightler, Foundations of optimization (1967).
[3] G. S. G. Bevridge, R. S. Schechter, Optimization: theory and practice (1970).
[4] J. W. Blakenuse, S. H. Davis, Optimization techniques (1964).
[5] M. J. D. Powell, An efficient method for finding the minimum of a function of several variables without calculating derivatives, *Computer J.*, 7 (1964), 155—162.
[6] R. Fletcher., M. J. D. Powell, A rapidly convergent descent method for minimization, *Computer J.*, **6** (1963), 163—168.
[7] J. Kiefer, Sequential minimax search for a maximum, *Proc. Amer. Math. Soc.*, **4** (1953), 502—506.
[8] W. Sperdley, Optimization (1969).
[9] A. Lavi, T. P. Vogl, Recent advances in optimization techniques (1966).

1970 年后国内各地出版的优选法文献和推广优选法成果资料汇编已达数百种，内容丰富，因总数浩繁，此处不能一一收录。

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

参 考 文 献

[1] D. J. Wilde, Optimum seeking methods (1964).

[2] D. J. Wilde, C. S. Beightler,Foundations of optimization (1967).

[3] G. S. G. Bevridge, R. S. Schechter, Optimization: theory and practice (1970).

[4] J. W. Blakenuse, S. H. Davis, Optimization techniques (1964).

[5] M. J. D. Powell, An efficient method for finding the minimum of a function of several variables without calculating derivatives, *Computer J.*, 7 (1964), 155—162.

[6] R. Fletcher., M. J. D. Powell, A rapidly convergent descent method for minimization, *Computer J.*, **6** (1963), 163—168.

[7] J. Kiefer, Sequential minimax search for a maximum, *Proc. Amer. Math. Soc.*, **4** (1953), 502—506.

[8] W. Sperdley, Optimization (1969).

[9] A. Lavi, T. P. Vogl, Recent advances in optimization techniques (1966).

1970 年后国内各地出版的优选法文献和推广优选法成果资料汇编已达数百种，内容丰富，因总数浩繁，此处不能一一收录。

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

参 考 文 献

[1] D. J. Wilde, Optimum seeking methods (1964).

[2] D. J. Wilde, C. S. Beightler,Foundations of optimization (1967).

[3] G. S. G. Bevridge, R. S. Schechter, Optimization: theory and practice (1970).

[4] J. W. Blakenuse, S. H. Davis, Optimization techniques (1964).

[5] M. J. D. Powell, An efficient method for finding the minimum of a function of several variables without calculating derivatives, *Computer J.*, 7 (1964), 155—162.

[6] R. Fletcher., M. J. D. Powell, A rapidly convergent descent method for minimization, *Computer J.*, **6** (1963), 163—168.

[7] J. Kiefer, Sequential minimax search for a maximum, *Proc. Amer. Math. Soc.*, **4** (1953), 502—506.

[8] W. Sperdley, Optimization (1969).

[9] A. Lavi, T. P. Vogl, Recent advances in optimization techniques (1966).

1970 年后国内各地出版的优选法文献和推广优选法成果资料汇编已达数百种，内容丰富，因总数浩繁，此处不能一一收录。

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

参 考 文 献

[1] D. J. Wilde, Optimum seeking methods (1964).

[2] D. J. Wilde, C. S. Beightler,Foundations of optimization (1967).

[3] G. S. G. Bevridge, R. S. Schechter, Optimization: theory and practice (1970).

[4] J. W. Blakenuse, S. H. Davis, Optimization techniques (1964).

[5] M. J. D. Powell, An efficient method for finding the minimum of a function of several variables without calculating derivatives, *Computer J.*, 7 (1964), 155—162.

[6] R. Fletcher., M. J. D. Powell, A rapidly convergent descent method for minimization, *Computer J.*, **6** (1963), 163—168.

[7] J. Kiefer, Sequential minimax search for a maximum, *Proc. Amer. Math. Soc.*, **4** (1953), 502—506.

[8] W. Sperdley, Optimization (1969).

[9] A. Lavi, T. P. Vogl, Recent advances in optimization techniques (1966).

1970 年后国内各地出版的优选法文献和推广优选法成果资料汇编已达数百种，内容丰富，因总数浩繁，此处不能一一收录。

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

参 考 文 献

[1] D. J. Wilde, Optimum seeking methods (1964).

[2] D. J. Wilde, C. S. Beightler,Foundations of optimization (1967).

[3] G. S. G. Bevridge, R. S. Schechter, Optimization: theory and practice (1970).

[4] J. W. Blakenuse, S. H. Davis, Optimization techniques (1964).

[5] M. J. D. Powell, An efficient method for finding the minimum of a function of several variables without calculating derivatives, *Computer J.*, 7 (1964), 155—162.

[6] R. Fletcher., M. J. D. Powell, A rapidly convergent descent method for minimization, *Computer J.*, **6** (1963), 163—168.

[7] J. Kiefer, Sequential minimax search for a maximum, *Proc. Amer. Math. Soc.*, **4** (1953), 502—506.

[8] W. Sperdley, Optimization (1969).

[9] A. Lavi, T. P. Vogl, Recent advances in optimization techniques (1966).

1970 年后国内各地出版的优选法文献和推广优选法成果资料汇编已达数百种，内容丰富，因总数浩繁，此处不能一一收录。

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

参 考 文 献

[1] D. J. Wilde, Optimum seeking methods (1964).

[2] D. J. Wilde, C. S. Beightler, Foundations of optimization (1967).

[3] G. S. G. Bevridge, R. S. Schechter, Optimization: theory and practice (1970).

[4] J. W. Blakenuse, S. H. Davis, Optimization techniques (1964).

[5] M. J. D. Powell, An efficient method for finding the minimum of a function of several variables without calculating derivatives, *Computer J.*, 7 (1964), 155—162.

[6] R. Fletcher., M. J. D. Powell, A rapidly convergent descent method for minimization, *Computer J.*, **6** (1963), 163—168.

[7] J. Kiefer, Sequential minimax search for a maximum, *Proc. Amer. Math. Soc.*, **4** (1953), 502—506.

[8] W. Sperdley, Optimization (1969).

[9] A. Lavi, T. P. Vogl, Recent advances in optimization techniques (1966).

1970 年后国内各地出版的优选法文献和推广优选法成果资料汇编已达数百种，内容丰富，因总数浩繁，此处不能一一收录。

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

参 考 文 献

[1] D. J. Wilde, Optimum seeking methods (1964).

[2] D. J. Wilde, C. S. Beightler, Foundations of optimization (1967).

[3] G. S. G. Bevridge, R. S. Schechter, Optimization: theory and practice (1970).

[4] J. W. Blakenuse, S. H. Davis, Optimization techniques (1964).

[5] M. J. D. Powell, An efficient method for finding the minimum of a function of several variables without calculating derivatives, *Computer J.*, 7 (1964), 155—162.

[6] R. Fletcher., M. J. D. Powell, A rapidly convergent descent method for minimization, *Computer J.*, **6** (1963), 163—168.

[7] J. Kiefer, Sequential minimax search for a maximum, *Proc. Amer. Math. Soc.*, **4** (1953), 502—506.

[8] W. Sperdley, Optimization (1969).

[9] A. Lavi, T. P. Vogl, Recent advances in optimization techniques (1966).

1970 年后国内各地出版的优选法文献和推广优选法成果资料汇编已达数百种，内容丰富，因总数浩繁，此处不能一一收录。

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

参 考 文 献

[1] D. J. Wilde, Optimum seeking methods (1964).

[2] D. J. Wilde, C. S. Beightler, Foundations of optimization (1967).

[3] G. S. G. Bevridge, R. S. Schechter, Optimization: theory and practice (1970).

[4] J. W. Blakenuse, S. H. Davis, Optimization techniques (1964).

[5] M. J. D. Powell, An efficient method for finding the minimum of a function of several variables without calculating derivatives, *Computer J.*, 7 (1964), 155—162.

[6] R. Fletcher., M. J. D. Powell, A rapidly convergent descent method for minimization, *Computer J.*, **6** (1963), 163—168.

[7] J. Kiefer, Sequential minimax search for a maximum, *Proc. Amer. Math. Soc.*, **4** (1953), 502—506.

[8] W. Sperdley, Optimization (1969).

[9] A. Lavi, T. P. Vogl, Recent advances in optimization techniques (1966).

1970 年后国内各地出版的优选法文献和推广优选法成果资料汇编已达数百种，内容丰富，因总数浩繁，此处不能一一收录。

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

参 考 文 献

[1] D. J. Wilde, Optimum seeking methods (1964).
[2] D. J. Wilde, C. S. Beightler, Foundations of optimization (1967).
[3] G. S. G. Bevridge, R. S. Schechter, Optimization: theory and practice (1970).
[4] J. W. Blakenuse, S. H. Davis, Optimization techniques (1964).
[5] M. J. D. Powell, An efficient method for finding the minimum of a function of several variables without calculating derivatives, *Computer J.*, 7 (1964), 155—162.
[6] R. Fletcher., M. J. D. Powell, A rapidly convergent descent method for minimization, *Computer J.*, **6** (1963), 163—168.
[7] J. Kiefer, Sequential minimax search for a maximum, *Proc. Amer. Math. Soc.*, **4** (1953), 502—506.
[8] W. Sperdley, Optimization (1969).
[9] A. Lavi, T. P. Vogl, Recent advances in optimization techniques (1966).

1970 年后国内各地出版的优选法文献和推广优选法成果资料汇编已达数百种，内容丰富，因总数浩繁，此处不能一一收录。

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

参考文献

[1] D. J. Wilde, Optimum seeking methods (1964).

[2] D. J. Wilde, C. S. Beightler, Foundations of optimization (1967).

[3] G. S. G. Bevridge, R. S. Schechter, Optimization: theory and practice (1970).

[4] J. W. Blakenuse, S. H. Davis, Optimization techniques (1964).

[5] M. J. D. Powell, An efficient method for finding the minimum of a function of several variables without calculating derivatives, *Computer J.*, 7 (1964), 155—162.

[6] R. Fletcher., M. J. D. Powell, A rapidly convergent descent method for minimization, *Computer J.*, **6** (1963), 163—168.

[7] J. Kiefer, Sequential minimax search for a maximum, *Proc. Amer. Math. Soc.*, **4** (1953), 502—506.

[8] W. Sperdley, Optimization (1969).

[9] A. Lavi, T. P. Vogl, Recent advances in optimization techniques (1966).

1970 年后国内各地出版的优选法文献和推广优选法成果资料汇编已达数百种，内容丰富，因总数浩繁，此处不能一一收录。

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

参 考 文 献

[1] D. J. Wilde, Optimum seeking methods (1964).

[2] D. J. Wilde, C. S. Beightler, Foundations of optimization (1967).

[3] G. S. G. Bevridge, R. S. Schechter, Optimization: theory and practice (1970).

[4] J. W. Blakenuse, S. H. Davis, Optimization techniques (1964).

[5] M. J. D. Powell, An efficient method for finding the minimum of a function of several variables without calculating derivatives, *Computer J.*, 7 (1964), 155—162.

[6] R. Fletcher., M. J. D. Powell, A rapidly convergent descent method for minimization, *Computer J.*, **6** (1963), 163—168.

[7] J. Kiefer, Sequential minimax search for a maximum, *Proc. Amer. Math. Soc.*, **4** (1953), 502—506.

[8] W. Sperdley, Optimization (1969).

[9] A. Lavi, T. P. Vogl, Recent advances in optimization techniques (1966).

1970 年后国内各地出版的优选法文献和推广优选法成果资料汇编已达数百种，内容丰富，因总数浩繁，此处不能一一收录。

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn